MASTER'S THESIS

Expectancy-Value, Tracking and Gender as Predictors of Achievement and Aspiration in Mathematics: a Study of 15-Year-Old Students in the Netherlands.

Roosken, Erik

Award date:
2020

Expectancy-Value, Tracking and Gender as Predictors of Achievement and Aspiration in Mathematics: a Study of 15-Year-Old Students in the Netherlands.

Expectancy-Value, Tracking en Gender als Predictors van Wiskundige Prestaties en Aspiraties: een Onderzoek onder Vijftien Jaar Oude Leerlingen in Nederland.

Erik Roosken

Master Onderwijswetenschappen
Open Universiteit

Date: 12 March 2019
Name supervisor: Dr. Kate Man Xu

Table of contents

Samenvatting

Expectancy-Value, Tracking en Gender als Voorspellers van Wiskundige Prestaties en Aspiraties: een Studie van Vijftien Jaar Oude Leerlingen in Nederland.

Erik Roosken

Volgens de expectancy-value theorie (EVT) worden prestatie-gerelateerde keuzes beïnvloed door de relatieve waarde en de mate van kans op succes bij een specifieke taak. Deze studie maakt gebruik van EVT en zijn multiplicatieve eigenschappen om de samenhang tussen leerlingenmotivatie voor wiskunde en onderwijsresultaten en aspiraties te onderzoeken. Het onderzoek is relevant omdat er wereldwijd een tekort is aan beroeps- en academisch geschoold personeel binnen het STEM-werkveld (Science, Technology, Engineering, Mathematics). Het zoeken naar mogelijke oplossingen voor de groeiende problemen binnen dit werkveld in Nederland, gebaseerd op de onderzoeksresultaten is daarom belangrijk.

Omdat er nog weinig onderzoek gedaan is naar de invloed van tracking, gender en motivatie van leerlingen op academische prestaties en aspiraties en in welke mate motivatie van leerlingen varieert in de verschillende tracks en seksen in de Nederlandse context, is dit de belangrijkste bijdrage van dit onderzoek. Het onderzoeken van genderverschillen in motivatie en prestaties en aspiraties is relevant omdat vrouwen ondergerepresenteerd zijn in het STEM-werkveld.

Het onderzoek is gebaseerd op secundaire data, doordat het PISA-data uit 2012 gebruikt. Het samplingontwerp is een ‘twee-niveaus gestratificeerd sample ontwerp’. Voor het onderzoek is de verzamelde data van 4224 leerlingen van 175 verschillende scholen gebruikt. Daarom is een hiërarchische structuur binnen de data waarschijnlijk en wordt in het onderzoek multilevel regressie gebruikt om de hypothesen te testen. Multilevel regressie zorgt ervoor dat de aanname van onafhankelijkheid niet wordt geschonden en het structureert de data op hiërarchische wijze.

De onderzoeksdata is verzameld door het OECD door middel van een leerlingenenquête en de PISA wiskundetoets. Van de verzamelde data zijn voor dit onderzoek het geslacht en het gevolgde onderwijsniveau van de leerlingen van belang. Verder van belang zijn de constructen wiskunde interesse (interest value), wiskunde gebruikswaarde (utility value), wiskunde zelfvertrouwen (self-efficacy) en wiskundige aspiraties (aspirations). Om de wiskundige onderwijsresultaten te meten, zijn de resultaten van de PISA-wiskundetoets 2012 gebruikt.

De resultaten tonen aan dat leerlingenmotivatie structureel voorspellend zijn voor onderwijsresultaten en aspiraties en dat er verschillen zijn tussen de verschillende onderwijstypen. In sommige gevallen is leerlingenmotivatie vergelijkbaar in de verschillende onderwijsniveaus, zoals het effect van zelfvertrouwen op onderwijsresultaten, maar er zijn ook duidelijke verschillen gevonden, zoals de invloed van wiskunde interesse op aspiraties. Wat betreft gender zijn er veel overeenkomsten, maar zijn

ook verschillen geconstateerd, bijvoorbeeld de grote invloed van interesse op onderwijsresultaten bij meisjes en de negatieve invloed van zelfvertrouwen op aspiraties bij jongens. Het muliplicatieve effect van expectancy en value is over het algemeen negatief of afwezig; uitzondering is het effect van de interactie tussen zelfvertrouwen en gebruikswaarde op onderwijsresultaten. Een aantal interacties is verschillend in de diverse onderwijstypen. Er zijn echter geen verschillen gevonden tussen jongens en meisjes.

Een tekortkoming van de huidige studie is het niet in achting nemen van de variatie in samenstelling van de middelbare scholen in Nederland. Een andere tekortkoming is het nieuw geïntroduceerde ‘forced choice’ format waarmee aspiraties zijn gemeten. Om aspiraties bij alle leerlingen te verhogen en om de verschillen tussen prestaties te verkleinen, zou het verplicht stellen van wiskunde op ieder onderwijsniveau en in ieder leerprofiel een oplossing kunnen zijn. Als laatste zou het vergroten van de gebruikswaarde van wiskunde bij leerlingen een vliegwiel kunnen zijn bij het verhogen van aspiraties.

Keywords

expectancy-value, tracking, gender, wiskunde, onderwijsresultaten, aspiraties, PISA, multilevel regressie

Summary

Expectancy-value, tracking and gender as predictors of achievement and aspiration in mathematics: a study of 15-year-old students in the Netherlands.

Erik Roosken

According to expectancy-value theory (EVT) achievement-related choices are influenced by the relative value and probability of success of a specific task. This study draws on EVT and its multiplicative character to explore the relations between student motivation for mathematics and educational outcomes. This study is relevant because there currently is a worldwide shortage of a vocationally and academically trained workforce in the STEM occupations. Therefore, finding possible solutions for the growing problems in the STEM field in the Netherlands, based on the results of the study, is important.

Because little research has yet been conducted on the subject, the main contribution of this study is to investigate the effect of student motivation on academic achievement and aspirations, and how the effect varies across tracks and genders. Exploring gender differences in motivational behaviour and mathematical achievement and aspirations is relevant because of the low number of students and in particular female students aspiring careers in the STEM field.

The research is based on a secondary dataset as it utilizes the PISA data of 2012. The sampling design used for the PISA assessment was a two-stage stratified sample design. For this study, data from 4224 students in 175 different schools is used. Hence, a hierarchical structure within the data is expected and therefore this study utilizes multilevel regression analyses for testing the hypotheses, as it ensures that the assumption of statistical independence is not violated due to the data being nested in a hierarchical fashion.

The data was collected by OECD using a student questionnaire and the PISA mathematics test. From the data obtained from the questionnaire, this study was interested in the gender of the students and in which track the students were enrolled. Furthermore, mathematics interest/enjoyment and mathematics instrumental motivations were utilized as interest and utility value. Mathematics self-efficacy was the expectancy component, while mathematics intentions/aspirations and student achievement, the results from the PISA 2012 mathematics test, were utilized as outcome variables.

The results demonstrate that student motivation was consistently predictive of achievement and aspirations but there was variation of the effect across tracks. The predictive role of student motivation was comparable in different tracks in some cases, such as the effect of self-efficacy on achievement, but different in other cases, such as the effect of interest value on aspirations. Although students motivation was comparable for both genders, there were some differences. For girls interest value had a clear effect on achievement. Self-efficacy had a negative effect on aspirations for boys, while no effect was detected for girls. The multiplicative effect of expectancy and value was largely negative or absent; only the interaction between self-efficacy and utility positively affected math achievement. Some interactions varied across tracks, however no variation across gender was detected.

Some limitations of the study include not have taken into account the variation in the composition of tracking in secondary schools in the Netherlands and the newly introduced ‘forced choice’ format for measuring aspirations used in the PISA test. A practical implication to increase math aspirations and reduce differences in math achievement, making mathematics obligatory in all learning profiles in all tracks is a possible solution. Furthermore, the positive effect of utility value implies that it may serve as an effective intervention target in order to increase math aspirations.

Keywords

expectancy-value, tracking, gender, mathematics, achievement, aspirations, PISA, multilevel regression

1 Introduction

1.1 Purpose of the study

Tracking on the grounds of academic achievement is common in most industrialized countries around the world. Also in the Netherlands where students, compared to most countries, are tracked in a large number of tracks at an early age when they embark in secondary education (Korthals & Dronkers, 2016). The Dutch education system utilizes explicit tracking, which is a highly visible form of tracking based on prior achievement where status differences between the tracks are clear to students, parents, peers and teachers (Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). The effects of tracking are anything but unambiguous and therefore subject to much heated debate (Lucas, 1999).

The effect of tracking on academic achievement and educational and career aspirations, and in particular on student motivation is unclear and little research has yet been conducted. The main goal and contribution of this study is to investigate how expectancy and value, drawing on the expectancy-value theory (EVT), predict achievement and career aspirations for 15-year-old students in the Netherlands across different tracks. In addition, gender differences in motivational behaviour and mathematical achievement and aspirations are explored using Dutch PISA data from 2012.

Thus, the main research question of this thesis is: *To what extent do gender and track-specific differences exist in student achievement and educational and career aspirations in mathematics, and how does the predictive effect of motivation on achievement and aspirations in mathematics vary across tracks and gender?* In order to thoroughly answer the main research question, it is broken down in four research questions: First, *Is there a main effect of motivation, tracking, and gender on math achievement and aspirations?* Then, *Is there a multiplicative effect between self-efficacy and interest value and between self-efficacy and utility value on math achievement and aspirations?* Followed by *Is there an interaction effect between tracking and motivation on math achievement and aspirations?* Lastly, *Is there an interaction effect between gender and motivation on math achievement and aspirations?*

The mathematical context in which this study is conducted is relevant, due to the growing shortages of a vocationally and academically trained workforce in the STEM occupations (Science, Technology, Engineering, and Mathematics) worldwide (Watt, Shapka, Morris, Durik, Keating, & Eccles, 2012), and thus also in the Netherlands. In 2016 there were 24.000 vacancies in STEM occupations in the Netherlands.[1] At the same time the number of students choosing STEM careers in the Netherlands is declining. An important contributor to the shortage of workforce in the STEM occupations is the low number of female students aspiring careers in the STEM field (Wang & Degol, 2017).

[1] See for full article on the 'Nationaal Techniekpact 2017' https://www.kivi.nl/nieuws/artikel/minister-bussemaker-ocw-ontvangt-voortgangsrapportage-en-monitor-van-nationaal-techniekpact

[2] OECD (2014) reported that the share of women in 2012 in the field of science was 41%, and that women represented 28% of the total workforce in the field of engineering, manufacturing, and construction. Due to the lack of vocationally and academically trained students choosing STEM careers, and the existing gender gap in STEM occupations, it is important to investigate how motivational behaviour varies by both track and gender in order to find possible solutions for these growing problems in the STEM field.

1.2 Theoretical framework

1.2.1 Expectancy-value theory.

Since Atkinson's influential study in 1957, the Expectancy-value Theory (EVT) of motivation has been a popular model explaining achievement, effort, and choice-related behaviour. Atkinson defined motivation as the belief of an individual about how well he or she would perform on an activity (expectancy) and to what extent he or she valued that activity (value) (Eccles & Wigfield, 2000; Eccles & Wigfield, 2002). Atkinson's study had however taken place in an experimental environment, therewith substantially lowering its ecological validity (Trautwein, Marsh, Nagengast, Lüdtke, Nagy, & Jonkmann, 2012; Nagengast, Marsh, Scalas, Xu, Hau, & Trautwein, 2011). In the last two decades of the twentieth century, Eccles and Wigfield (2002) contributed significantly to EVT by applying the theory to real-world situations (Trautwein et al., 2012). They also extended EVT by differentiating the value component into four dimensions: interest, attainment, utility, and cost. In recent decades EVT has been widely used to explain performance, persistence and task choice (Eccles & Wigfield, 2002).

EVT is thus composed of an expectancy component and a value component. Both expectancies and values are assumed to be influenced by a large array of factors, such as the cultural milieu of the child, the child's goals and self-schemata, and previous achievement-related experiences (Eccles & Wigfield, 2002). See figure 1 for a detailed representation of the Eccles et al. expectancy-value model of achievement. Furthermore, research indicates that associations between expectancy and value beliefs increase within a specific domain as students grow older (Wigfield, Eccles, Yoon, Harold, Arbreton, Freedman-Doan, & Blumenfeld, 1997).

---

[2] See https://www.nu.nl/carriere/5244989/technische-sector-verwacht-concurrentieproblemen-personeelstekort.html for the full article on the problem of growing shortages in the STEM occupations and the lack of women choosing STEM careers.

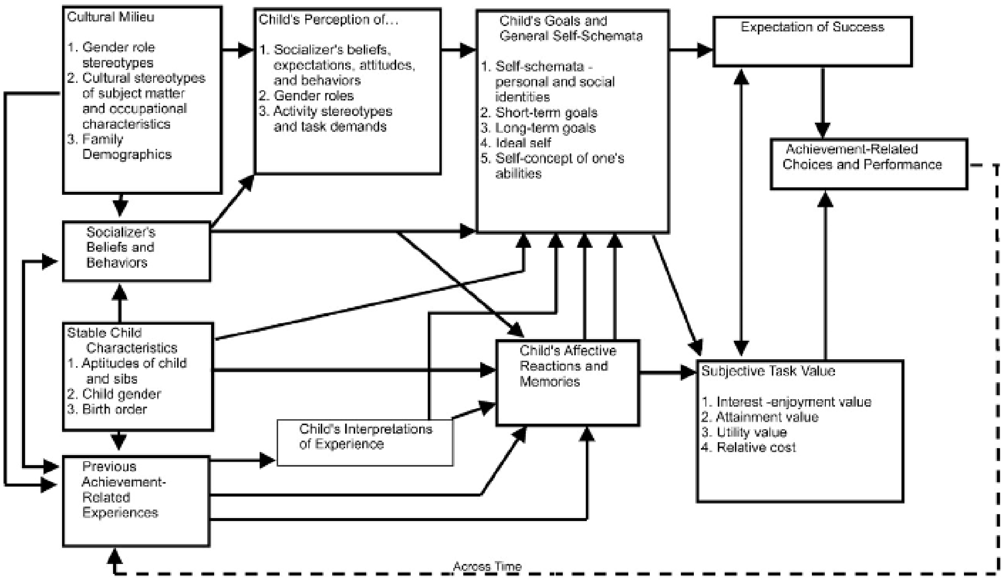

Figure 1. The Eccles et al. expectancy-value model of achievement (Eccles & Wigfield, 2002, p. 119.)

The expectancy component can be defined as a task-specific belief about success in an upcoming academic activity or as an evaluation of an individuals' competence in a particular area in either the immediate or longer-term future (Eccles & Wigfield, 2002). Both definitions of expectancy are highly correlated and are often used interchangeably. Expectancy is closely linked to other conceptions of self-belief, such as academic self-concept and self-efficacy (Trautwein et al., 2012). The present study uses self-efficacy as the expectancy component.

Self-efficacy refers to convictions about one's learning abilities. An abundance of research has demonstrated that self-efficacy influences academic motivation, learning, and achievement. Self-efficacy originated in Bandura's social cognitive theory (1986, 1997), which presumes that human achievement depends on interactions between an individual's behaviours, his thoughts and beliefs, and his surroundings. Learners learn to assess their self-efficacy from their own achievement, observations of others, praise and encouragements from their environment, and their physiological reactions. Self-efficacy beliefs influence task choice, effort, persistence, resilience, and achievement (Bandura, 1997; Schunk, 1995). Therefore, students with high self-efficacy participate more readily, work harder, persist longer when faced with difficulties, and achieve at a higher level (Pajares & Schunk, 2001).

Although the link between self-efficacy and achievement is stronger than the link with aspirations (Wang, 2012), academic self-efficacy has also been found to be an important predictor for career decisions at the individual level. Betz and Hackett (1981), for instance, demonstrated that women's preference for traditionally female occupations could not be explained by differences in math achievement but could be related to their low self-beliefs in relation to math invested careers.

The other constituent of EVT is value. Value beliefs are firm predictors of choice, effort and persistence (Trautwein et al., 2012), and have been differentiated into four separate dimensions by Eccles and Wigfield (2002). The four dimensions being: intrinsic value, utility value, attainment value, and cost (Eccles & Wigfield, 2002). Intrinsic value (i.e. interest motivation) is defined as the amount of pleasure which an individual gains while performing a task or his or her personal interest in a particular subject. In forming interest in a particular activity parents, peers, and teachers are highly important. Related to mathematics, interest value means that a student enjoys doing a mathematics activity and/or likes mathematics in general. Interest is a crucial factor determining the quality of learning as well as educational and career choices and aspirations. Interest values seldomly affect achievement. Furthermore, math interest is found to decrease when students get older and social interests increase, as do task complexity and demands for effort (Nagy, Trautwein, Baumert, Köller, & Garrett, 2006). Moreover, Frenzel, Goetz, Pekrun, and Watt (2010) found that students in Germany in lower tracks had more favourable interest trajectories for mathematics than students in higher tracks.

Utility value is defined as the perceived future usefulness of engagement and achievement in activities in a specific domain. Future goals can be career aspirations or gaining access to a desired school. Related to mathematics, utility value is the degree to which see mathematics as essential for attaining a desired future goal. Utility value is thus not necessarily related to interest value (Eccles & Wigfield, 2002). Research by Guo, Marsh, Parker, Morin, & Yeung (2015) demonstrated that utility value has a significant effect on educational aspirations but only a weak effect on achievement.

Attainment value is the personal importance of doing well on a task and is also related to the relevance of engaging in a task in order to confirm or disconfirm one's self-schema in task-specific domains. (Eccles & Wigfield, 2002). Lastly, Cost can be defined as the negative consequences of engaging in a specific task, such as time, effort and feelings of anxiety. Linked to mathematics cost refers to students calculating how much time and effort it will take to engage in a specific task or whether they are willing to experience feelings of fear or anxiety while performing a task (Eccles & Wigfield, 2002; Trautwein et al., 2012). The present study utilizes interest and utility value as the value components of EVT.

The cornerstone of classic EVT developed by Atkinson was that achievement-related behaviour was a multiplicative function of expectancy and value. Effectively, when an individual placed more value on a specific task, the effect of expectancy on motivation would be even stronger. With the embarkment of modern EVT however, including the highly influential article on modern EVT by Wigfield & Eccles (2000), the multiplicative character disappeared. From then on studies treated EVT as an additive model rather than a multiplicative model (Nagengast et al., 2011). The multiplicative or interactive model suggests that the combined effect of expectancy and value beliefs differs from the sum of the two separate effects, because both predictors affect the criterion variable in the same direction. Therefore they produce a stronger effect on outcomes than an additive model (Trautwein et al., 2012). Nagengast et al. (2011) demonstrated that the multiplicative model of EVT had unjustly disappeared and was still relevant in modern research on achievement motivation. A statement validated by Trautwein et al. (2012) and Guo et al. (2015). So, while additive models are compensatory, multiplicative models are not; both expectancy as well as value beliefs must be high in order to instigate motivated behaviour (Nagengast et al., 2011). An important contribution of this study is testing the multiplicative effect of expectancy and value in predicting achievement and aspirations in the Dutch context.

1.2.2 Tracking.

Tracking is the separation of students into tracks that vary in academic orientation and curricula, mostly based on prior achievement, and is widely used in secondary schools around the world. Tracking can be implemented in various ways. Differences between countries are often the amount of tracks offered to students, usually ranging from two to five, and the age of students when tracked, typically between ten and sixteen (Dockx, De Fraine, & Vandecandelaere, 2018; Korthals & Dronkers, 2016).

Two types of tracking are distinguished. Explicit school-level tracking is a highly visible form of tracking students into different tracks often in different schools or buildings, based on prior achievement where status differences between the tracks are clear to students, parents, peers and teachers. Explicit tracking thus also leads to tracks with different teachers often in different buildings, thereby imposing peer homogeneity on students within the same track (Korthals & Dronkers, 2016; Maaz, Trautwein, Lüdtke, & Baumert, 2008). Implicit school-level tracking is tracking on the basis of factors such as area of residence (Trautwein et al., 2006). In countries which do not stream explicitly, non-institutionalized tracking can occur through ability grouping (i.e. different classes in the same school) or seating (i.e. different curricula within classes) (Korthals & Dronkers, 2016; Maaz et al., 2008; Ireson, Hallam, & Hurley, 2005).

Because tracking is so widely utilized across educational systems, the question why students are tracked is fair to ask. The main goal of tracking is to create learning environments tailored to different groups of students. Student groups are often based on academic ability and therefore homogenous. Homogenous groups create the possibility to focus curricula and teachers on specific learning needs, which benefits academic performance and skill specialization. However, by creating achievement based groups, tracking also institutionalizes social distance between students groups (Dockx et al., 2018).

1.2.2.1 Achievement and aspirations.

Many researchers have debated over the effect of tracking on students achievement, future educational careers, morale, and happiness (Trautwein et al., 2006; Maaz et al., 2008). However, studies show ambiguous results and there is much debate on the effects of tracking on achievement and aspirations of students (Trautwein et al, 2006; Ireson et al., 2005).

Critics of tracking argue that students in vocational tracks are at a disadvantage compared to students in tracks that prepare for higher and academic education. Students in lower tracks receive lower quality teaching which could lead to lower achievement (Trautwein et al., 2006; Maaz et al., 2008). Past research has almost consistently found that a higher tracks positively affects academic achievement, while the opposite is true for lower tracks. Studies conducted over time and in different national contexts with various ways of tracking, controlling for initial ability and other student characteristics, have confirmed that students achieve more success in higher tracks. This effect is the most evident for mathematics achievement (Guill, Lüdtke, & Köller, 2017; Van Houtte, 2017). Furthermore, Maaz et al. (2008) came to conclusion that explicit tracking contributes to strengthening the link between socioeconomic background and student achievement, while Becker et al. (2012) even found an increasing gap in psychometric intelligence (i.e. IQ) between academic-tracked students and students of vocational tracks in Germany.[3] Very recent research, performed by Dockx et al. (2018) in the Flemish context, also showed that tracking along cognitive lines was beneficial for students in higher tracks in particular for reading and comprehension and mathematics.

Korthals (2015) however, posed that tracking works favourably for all students when they are tracked strictly by academic achievement and at an early age, as teachers are more able to align the curriculum to a homogenous group of students. Korthals demonstrated that achievement of students in reading, science, and mathematics were favourably affected by early tracking in a high number of tracks.

Concerning tracking and aspirations, there is sufficient evidence pointing to negative effects of tracking on student's educational and career aspirations (Trautwein et al., 2006; Maaz et al., 2008; Dupriez, Monseur, Van Campenhoudt, & Lafontaine, 2012). Students in lower tracks develop lower mathematical aspirations than students in academic-oriented tracks. Academic tracks prepare students for higher education, but this is not the first goal of vocational tracks. Thus, vocational students are less likely to aspire to higher education and therefore tend to be less engaged in school. Students who are not planning for higher education generally perceive that grades do not matter much for their career aspirations (Wang and Eccles, 2012; Van Houtte, 2017).

[3] The IQ was measured using the Figure Analogies subscale of the *Kognitiver Fähigkeitstest*, which is a slightly adapted German version of Thorndike's Cognitive Abilities Test.

In a cross-national study, Dupriez et al. (2012) found that in countries where tracking occurs before the age of 15, the mathematical aspirations of students in academic tracks were higher than students in lower tracks. Remarkably, in the Netherlands, amongst other countries as well, the mean score of the socioeconomic background of the school was a more significant predictor for mathematical aspirations than the school's mean mathematical score. Although remarkable, it is not a surprise: in the Netherlands socioeconomic background is traditionally a firm predictor of student's track enrolment (Inspectie van Onderwijs, 2016).

Due to the effect of tracking on achievement and educational and career aspirations, this study will use tracking as a predictor and a moderator. First of all, the differences in mathematical achievement and aspirations are analysed. Secondly, the study explores to what extent tracking moderates motivational behaviour in predicting achievement and aspirations.

1.2.2.2 Tracking in the Netherlands.

In the Dutch education system explicit tracking is common. Until 2015 all pupils took an obligatory exit test in their last primary school year which in the majority of the cases determined which track they enrolled in secondary school. The data used in this study is from 2012 when students were still tracked according to a standardized national test taken by all sixth-graders in the Netherlands. Although admission procedures changed in 2015, it is still important to investigate the effect of tracking in the older system in order to better evaluate the effect of the new system in comparison to previous practises. Moreover, the focus on mathematical motivation and career aspirations is unique to PISA 2012.

In the Netherlands students are tracked at an early age (i.e. around 12 years old). Research by Werfhorst and Mijs (2010) concluded that early tracking enhances the effect of social economic background on educational achievement. According to the critics tracking thus leads to growing educational inequity which according to the PISA-data of 2015 is a growing problem in the Netherlands: Origin and social economic background still largely decide educational achievement. Furthermore, research on the PISA data of 2009 showed that tracking at a later age generally enhances equity and achievement Lavrijsen, Nicaise, & Wouters, 2013).

The number of tracks in the Dutch educational system is high compared to other countries. Korthals and Dronkers (2016) concluded that when schools take prior achievement into account, an increasing number of tracks benefits student achievement in reading, mathematics and science. The Dutch secondary education consists of three main tracks which prepare students for different academic and professional futures. Voorbereidend Middelbaar Beroepsonderwijs or VMBO (preparatory applied/vocational education), Hoger Algemeen Voorbereidend Onderwijs or HAVO (preparatory higher professional education) and Voorbereidend Wetenschappelijk Onderwijs or VWO (preparatory university level education). These tracks do not only differ in their curricula, but also in their duration and future possibilities.[4]

VMBO type of education lasts a total of four years. Within the VMBO track there are four different learning tracks: VMBO-Basisberoepsgerichte Leerweg (VMBO-BL), the basic vocational track; VMBO-Kaderberoepsgerichte Leerweg (VMBO-KL), the advanced vocational track; VMBO-Gemengde Leerweg (VMBO-GL), combined track of theoretical courses and practical subjects; and finally VMBO-Theoretische Leerweg (VMBO-TL), the theoretical track. The different learning tracks within VMBO range from practical to theoretical education, but also increased cognitive ability needed to conclude the learning track. When a pupil receives a VMBO advice when concluding primary education, it is always specified to one of these four learning tracks. Furthermore, students who conclude either the combined or the theoretical track have more educational and thus future professional possibilities, including admission to the HAVO track, than students concluding the basic or advanced vocational track. Because VMBO-GL and VMBO-TL require the same cognitive abilities of students, they form one track during the analysis.

VMBO starts with a general phase of two years. During these two years students are offered a broad range of subjects (i.e. sciences, mathematics, Dutch, English, and French or German language). Starting the third year, students choose a particular learning profile they would like to further study. In VMBO mathematics is obligatory in six of the ten profiles. These are the technical oriented profiles and the profile preparing for agricultural professions. In the other four profiles mathematics is an elective course.

HAVO lasts five years, while VWO lasts six years. HAVO secondary education prepares students for admission at a four-year applied university level education. Finishing the applied university level, or higher vocational education as it is called in the Netherlands grants students a bachelor’s degree. VWO secondary education prepares students for admission at university level, granting students a master’s degree. After a general phase of three years, students in the HAVO and VWO tracks choose a learning profile which partly determines future career paths. In VWO mathematics is obligatory in all profiles. In HAVO mathematics is obligatory in three of the four different profiles.

Because each track prepares students for different professional careers, this study is thus concerned with different levels of achievement, aspirations, interest and utility value, and self-efficacy. Although cross-track achievement comparisons have long been explored by national Dutch organiza-

---

[4] Although I regard the different tracks in Dutch secondary system as a continuum rather than a hierarchical system ranging from vocational tracks to academic oriented tracks, I will for practical matters refer to ‘higher’ and ‘lower’ tracks, with the academic track VWO as being the ‘highest’ track.

tions (i.e. CITO), findings of the present study will add another layer of understanding when comparing the relations between mathematical achievement, aspirations, and motivation across the various tracks drawing on EVT.

1.2.3 Gender.

1.2.3.1 Achievement.

This study also focuses on gender differences in mathematics achievement and aspirations. Recent research literature indicates that the gender difference in math achievement has narrowed. Whenever gender differences are found, reasons are often diverse and elusive (Gallagher & Kaufman, 2005). Hyde (2005) proposed the Gender Similarities Hypothesis which stated that males and females share most psychological traits and differ on some. In her meta-analysis, Hyde concluded that in mathematics the Gender Similarities Hypothesis was verified, although differences were found in computation favouring girls in elementary and middle school and in complex problem solving in high school favouring boys. Effect sizes however were small.

In another meta-analysis performed by Else-Quest, Hyde, & Linn (2010) based on TIMSS (Trends in International Mathematics and Science Study) and PISA data of 2003 the achievement gap had almost diminished in the overall sample. The weighted mean effect size of gender difference in achievement of the TIMSS was $d = -0{,}01$, with 63% of the effect sizes measured in 47 countries was below $d = 0{,}10$, meaning the effect size is negligible. Gender difference in the Netherlands in math achievement was $d = 0{,}10$, thus favouring boys. Concerning the PISA-results, there was a small gender difference in the overall sample in math achievement ($d = 0{,}11$). For the Netherlands, the gender difference in math achievement was smaller than the overall sample ($d = 0{,}06$). (Else-Quest et al., 2010). The results for both the TIMSS and PISA-data demonstrated that the gender difference in math achievement favouring boys had almost diminished in the Netherlands in 2003.

In a study by Forgasz and Leder (2017) the 2003-2006-2009-2012 PISA-data from Australia, Canada and the UK were analysed and compared. Forgasz and Leder did find a consistent pattern of higher male achievement in all three countries. In Canada the gap decreased over time, while in Australia the gap seemed to be widening. In the UK the picture was inconsistent. Their analyses did not include statistical procedures to test whether the gender differences were significant.

1.2.3.2 Aspirations.

Past research on gender differences in educational and career aspirations provides a mixed picture: in some studies males are reported to have higher aspirations, in other studies females have higher aspirations, and there are also studies which have found no significant gender differences. Yet, the dominant picture has been that male students are more likely to aspire to math-related careers, while female-students more often aspire careers which involve interaction with people, helping others and careers which are socially meaningful or important (Watt et al., 2012). Research by Wang and Degol (2017) has also demonstrated that women prefer working with people, while men prefer working with objects.

Recent research by Lauermann, Tsai, & Eccles (2017) confirmed the picture of male students still having higher mathematical aspirations than female students. Also when it comes to the prestige dimension of occupations, STEM (Science, Technology, Engineering, and Mathematics) careers are often held as highly prestigious and are dominated by males. Females more often choose prestigious occupations which are less mathematically intensive, such as becoming a lawyer (Watt et al., 2012; Riegle-Crumb, Moore, & Ramos-Wada, 2011). An important explanatory factor for the gap in mathematical aspirations between males and females are the persisting gender stereotypes and implicit bias by teachers as well as parents (Wang & Degol, 2017; Forgasz & Leder, 2017).

Due to the effect of gender on achievement and educational and career aspirations, this study will use gender as both a predictor as well as a moderator. First of all, it is investigated whether the gender gap in mathematical achievement and aspirations still exists in the Netherlands. Secondly, the study explores whether gender moderates motivational behaviour in predicting achievement and aspirations.

1.3 Research questions and hypotheses

As stated in the introduction, the main research question in this study is: *To what extent do gender and track-specific differences exist in student achievement and educational and career aspirations in mathematics, and how does the predictive effect of motivation on achievement and aspirations in mathematics vary across gender and tracks**?*** This study aims to clarify these uncertainties by exploring the following research questions, which are visually represented in figure 1.2.

1.3.1 Research question 1.

Is there a main effect of motivation, tracking, and gender on math achievement and aspirations? First, the effects of tracking and gender are analysed. Based on previous research it is predicted that students in higher tracks score higher on achievement and aspirations than students in lower tracks. Concerning gender, there is growing evidence that achievement is independent of gender. It is therefore expected that there are no significant differences between achievement scores between boys and girls. Concerning aspirations it is expected that male students have higher math-related aspirations than girls.

Then, the main effects of self-efficacy, interest value and utility value, controlling for gender and tracking, on achievement and aspirations are analysed. An abundance of research has demonstrated the positive effects of self-efficacy on achievement and aspirations. Interest and utility value are expected to predict aspirations only.

1.3.2 Research question 2.

Is there a multiplicative effect between self-efficacy and interest value and between self-efficacy and utility value on math achievement and aspirations? The multiplicative character of the expectancy-value theory is explored by analysing the interaction effect of expectancy and value variables, while controlling for gender, tracking and main effects of motivation. Firstly, the interaction between self-efficacy and interest value is analysed. Thereafter, the interaction between self-efficacy and utility value is investigated. It is predicted that all multiplicative effects are significant predictors of achievement and aspirations.

1.3.3 Research question 3.

Is there an interaction effect between tracking and motivation on math achievement and aspirations? First, the main effects of gender and motivation on math achievement and aspirations within the five tracks are analysed. Secondly, the interactions between the tracks and motivation are analysed. Then, the multiplicative effects of expectancy and value within the five tracks are examined. Lastly, the interactions between the tracks and expectancy-value are analysed.

It is predicted that the interaction between tracking and self-efficacy is significant for achievement in all tracks. The interaction of tracking and utility value is expected to significantly predict math-related aspirations with a stronger effect in higher tracks. The interaction of tracking and interest value is predicted to be significant for aspirations, particularly in lower tracks, since there is growing evidence of positive interest trajectories in lower tracks.

1.3.4 Research question 4.

Is there an interaction effect between gender and motivation on math achievement and aspirations? First, the main effects of tracking and motivation on math achievement and aspirations for female and male students are examined. Secondly, the interactions between gender and motivation are analysed. Then, the multiplicative effects of expectancy-value for female and male students are examined. Lastly, the interactions between gender and the multiplicative effects of expectancy and value are scrutinized.

It is expected that the interaction of gender and self-efficacy has a significant effect on achievement and aspirations. Although male students generally have higher self-efficacy, female students with similar scores for self-efficacy tend to have higher achievement scores. Thus, it is expected that the interaction will show a stronger effect of self-efficacy on achievement for girls than for boys. It is also expected that the interaction of gender and utility value has a significant effect on aspirations, in particular for male students. The interaction effect of gender and interest value is predicted to significantly predict math-related aspirations, particularly for boys.

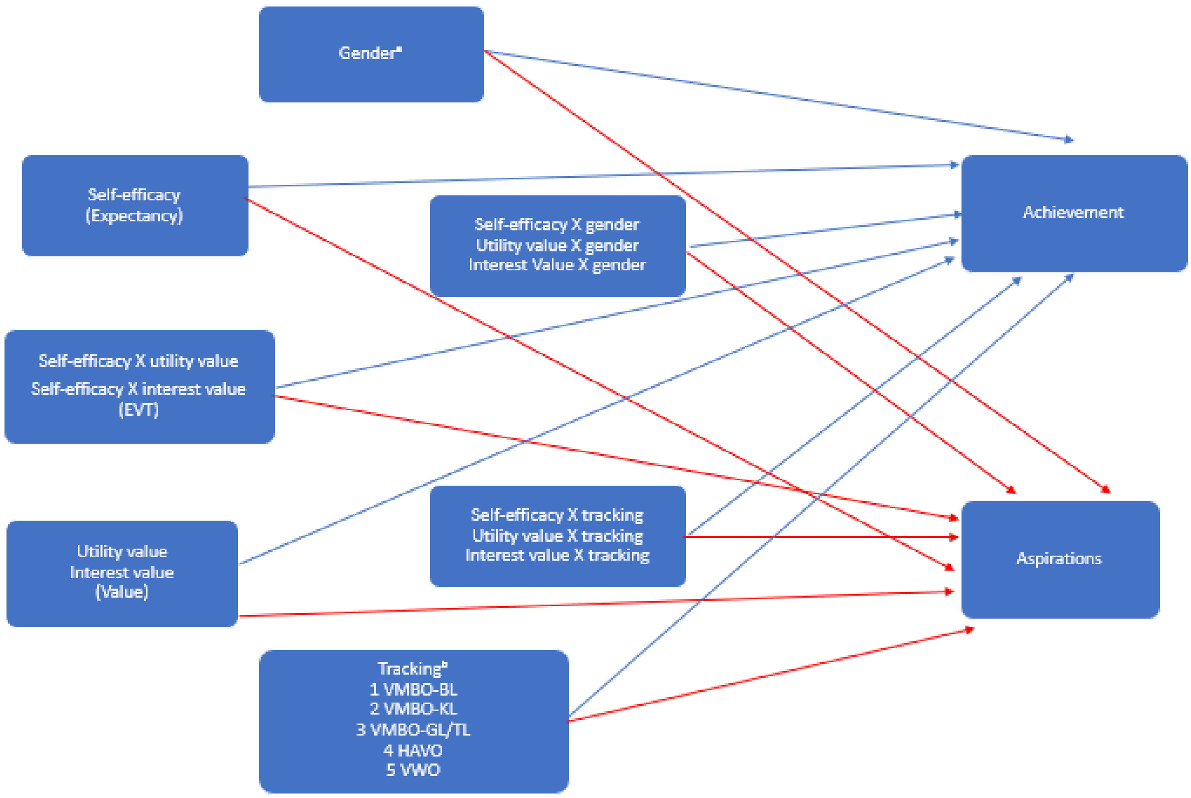

Figure 2. Visual representation of the research questions.
[a] When analysing the interaction effect of gender and motivation, the female students will be used as the reference category. The male students will be represented by a dummy variable.
[b] When analysing the interaction effect of tracking and motivation, the VMBO-BL track will be used as the reference category. Four dummy variables will represent each of the remaining tracks.

2 Method

2.1 Design

This research is cross-sectional, based on a secondary dataset as it utilizes the database of the Programme of International Student Assessment (PISA) database of 2012 to answer the abovementioned research questions and hypotheses. The PISA is a triennial international survey which commenced in 2000, aiming to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students. It consists of an array of tests (reading, mathematics, and science) and questionnaires developed by The Organization for Economic Cooperation and Development (OECD). The PISA strives to investigate how well students are prepared to meet the challenges of the future, rather than how well they master particular curricula. The internationally comparable information provided by PISA allows countries to assess how well their 15-year-old students are prepared for life in a larger context and to compare their relative strengths and weaknesses (OECD, 2012).

2.2 Sample

In this study the PISA database of 2012 will be utilized, which was the programme's 5th survey. It assessed the competencies of 15-year-olds in reading, science and a particular focus on mathematics in 65 countries. Approximately 510.000 students participated in PISA 2012, representing about 28 million 15-year-olds globally (Technical Rapport, 2012). This research is solely concerned with the data gathered in the Netherlands. Of the total of 194.000 Dutch 15-year-olds, the target population, 4460 students from 177 different schools, the actual sample, participated in the PISA survey. 127 of the students were enrolled in practical education and 109 students were in their first or second year of the vocational track. For practical matters these students are not taken into account during the analysis. Of the remaining 4224 students from 175 different schools, 2058 (48,7%) of them are female and 2166 male (51,3%). 396 students were enrolled in VMBO-BL, of which 179 (45,2%) were female and 217 (54,8%) male. 555 students received education on the level of VMBO-KL, of which 246 (44,3%) were female and 309 (55,7%) male. 1172 students were enrolled in either VMBO-GL or VMBO-TL. Of which 558 (47,6%) were female and 614 (52,4%) male. A total of 1088 students were enrolled in the track preparing for higher professional education, HAVO. 539 (49,5%) of them were female and 549 (50,5%) were male. In VWO, the track preparing for academic education, there were 536 (52,9%) female students and 477 (47,1%) male students, making a total of 1013 students enrolled in VWO.

2.3 Measures

The 2012 PISA data was collected by OECD using a student questionnaire and the PISA mathematics test. The PISA 2012 Student Questionnaire (SQ) consisted of six sections which questioned students about themselves, their family and home, their mathematic experiences, their school, their learning of mathematics, and their problem solving experiences. Of Section A, about the students themselves, this study is interested in the gender of the students and in which national study programme (i.e. track) the students are enrolled.

Of Section C four constructs are used. First of all mathematics interest/enjoyment, which is used to measure interest value. The PISA 2012 SQ assessed for mathematics interest using four survey items in order to measure the level of interest and enjoyment (INTMAT).[5] Students were asked to respond to four positive statements about mathematics, for example: "I enjoy reading about mathematics." Students were able to respond by indicating to what extent they agreed or disagreed with the statements. Each survey item was scored on a 4-point Likert scale ranging from (1) 'strongly agree' to (4) 'strongly disagree'. The reliability (Cronbach's alpha) of this construct was α=.86. The values of students' responses, for all constructs, were all inverted for Item Response Theory (IRT) scaling, with positive values indicating high levels of interest and negative values indicating low levels of such (OECD, 2014).

Secondly, mathematics instrumental motivations, for measuring utility value. The PISA 2012 SQ measured instrumental motivation (INSTMOT) using four survey items. Students were asked to respond to positive statements about mathematics, for example: "Making an effort in mathematics is worth it because it will help me in the work that I want to do later on." Students were able to respond by indicating to what extent they agreed or disagreed with the statements. Each survey item was scored on a 4-point Likert scale ranging from (1) 'strongly agree' to (4) 'strongly disagree'. The reliability (Cronbach's alpha) of this construct was α=.88 (OECD, 2014).

Thirdly, mathematics self-efficacy to measure student expectancy. The PISA 2012 SQ measured mathematics self-efficacy (MATHEFF) using eight survey items. The response categories were 'Very confident' (1), 'Confident' (2), 'Not very confident' (3), and 'Not at all confident' (4). For this construct, item difficulties ranged from a relatively easy one "Solving an equation like 3x+5= 17", to more difficult ones, such as "Finding the actual distance between two places on a map with a 1:10 000 scale" and "Calculating the petrol consumption rate of a car". The reliability (Cronbach's alpha) of this construct was α=.85 (OECD, 2014).

The last construct of Section C is mathematics intentions, which is used to measure student mathematical aspirations. The PISA 2012 SQ measured mathematical intentions (MATINTFC) using five items of the so-called 'Forced Choice' format which was one of the new item types employed in PISA 2012. This item type forced students to choose between mathematics and either language or science with respect to additional courses at school and beyond. The items forcing students to choose between mathematics and the test language were easier than the items that force students to choose between mathematics and science. The reliability (Cronbach's alpha) of this construct was α=.66 (OECD, 2014).

In order to measure student achievement, the PISA 2012 mathematics test is used. The mathematical ability of students was assessed through 56 mathematics units comprising a total of 110 cognitive items. The items were distributed across four mathematical content domains, namely *change and*

---

[5] All survey items of all the used constructs can be found in the Appendix.

relationships, *space and shape*, *quantity*, and *uncertainty and data*. Each content domain delivered approximately 25 percent of the questions for the test. The scores in the dataset used for statistical analyses do not represent the score attained by the students. Instead, five plausible values for each student are presented. The plausible values are random numbers drawn from the distribution of scores that could be reasonably assigned to each individual student. Plausible values as a set are better suited to describing the performance of the population, because the PISA test is an assessment of students' cognitive abilities, which is a construct not directly observed by the PISA test. Every student received five plausible values per domain and five plausible values for an overall mathematics score. This approach is based on the imputation theory of Rubin (1987). For this study the first plausible value was utilized.

2.4 Procedure

The sampling design used for the PISA assessment was a two-stage stratified sample design. The first-stage sampling units comprised of individual schools having 15-year-old students. Schools were sampled systematically from a national list of all PISA-eligible schools. Prior to sampling, schools in the sampling frame were assigned to mutually exclusive groups based on school characteristics called explicit strata, formed mainly to improve the precision of sample-based estimates. The second-stage sampling were students within sampled schools. Once schools were selected to be in the sample, a complete list of each sampled school's 15-year-old students was prepared. For each country a Target Cluster Size *(TCS)* was set, this value was typically 35 students. From each list of students that contained more than the *TCS*, a sample of typically 35 students were selected with equal probability and for lists of fewer than the *TCS*, all students on the list were selected (OECD, 2014).

2.5 Data-analysis

The data used for this study is collected amongst 4460 students from 177 different schools in the Netherlands. Hence, the presence of a hierarchical structure within the data is expected. The hierarchical structure stems from the fact that students are taught by different teachers in different schools. Also, answers given by the students are not independent from each other. This causes a violation of the statistical analyses assumption of independence (Field, 2013).

Within educational research it is necessary to take hierarchical structure into consideration, or else the correct interpretation of results is at risk. Therefore, research attempting to identify predictors of academic achievement is required to utilize a statistical analysis which respects the hierarchical nature of the data collected and the assumption of independence. This study thus makes use of multilevel regression analyses for testing the hypotheses. Multilevel regression analyses ensures that the assumption of independence is not violated. Furthermore, it nests the data in a hierarchical fashion during the regression analyses. Generally, the data is nested in a two-level hierarchy; level one being the students and level two the class or the school that the students belong to. Also, multilevel analyses, as opposed to regular regression analyses, gives the opportunity to use varying intercepts, slopes, or both. The present study uses multilevel regression analyses consisting of varying intercepts and fixed slopes, because it fits the data best. Models made up of varying intercepts and fixed slopes assume that the models for different variables are in different locations (i.e. different intercepts but have the same shape (Field, 2013). By using multilevel regression, school level differences are controlled for, which assures the interpretation of results are accurate and unbiased.

This study also employs moderation during the data-analyses. Moderation models are used to demonstrate the extent to which a variable (Z) affects the direction or strength of the relationship between an independent variable (X) and a dependent variable (Y). According to Baron and Kenny (1986) the moderator can affect the dependent variable in three ways, namely linear, quadratic and step. This research assumes the moderator has a linear effect. Baron and Kenny's method also poses a two-step evaluation of moderation effects. First, the effect of the independent variable on the dependent variable while controlling for the moderator variable is examined. Then, the effect of the interaction variable is tested when controlling for the main effects of the moderator and the independent variables (Baron & Kenny, 1986).

3 Results[6]

In this section the most important results of the data analysis are presented. As will be further expounded in the following paragraphs, the results demonstrate that student motivation is consistently predictive of achievement and aspirations, but there is variation of the effect across tracks. The predictive role of student motivation is comparable in different tracks in some cases, such as the effect of self-efficacy on achievement, but different in other cases, such as the effect of interest value on aspirations. Concerning gender, student motivation is comparable for boys and girls, but there are some differences. For girls interest value had a clear effect on achievement. Self-efficacy had a negative effect on aspirations for boys, while no effect was detected for girls. The multiplicative effect of expectancy and value was largely negative or absent; only the interaction between self-efficacy and utility positively affected math achievement. Some interactions varied across tracks, however no variation across gender was detected.

3.1 Descriptive statistics

Descriptive statistics for mathematic self-efficacy, interest value, utility value, achievement, and aspirations are presented in table B1. First, metric of the variables for the total sample are values in the

[6] All tables and figures form the Result section are found in the appendix.

original scale (factor scores issued by PISA; see relevant section in Method). Then, track differences and gender differences are presented using standardized metric values.

There were mean differences in motivational variables across students in different tracks. Students in in the VMBO-BL track scored lowest for self-efficacy (-.61), while VWO students had the highest scores (.52). The scores for interest value revealed similar patterns (scores ranging from -.20 to .11), with the exception that VMBO-KL students had the lowest score. The scores for utility value also showed increasing scores in higher tracks (scores between and -.17 and .15). Similar for math achievement, scores ranged from -1.36 to 1.02 from low to high tracks. The trend for math aspiration was somewhat different compared to the other variables: VMBO-BL students had the lowest aspirations (-.15), while HAVO students had the highest aspirations (.09). VMBO-KL students scored higher (-.01) than VMBO-GL/TL students (-.07). The difference between the lowest and the highest score was the smallest among all variables.

Regarding gender, male students scored higher for self-efficacy (.21 compared to -.22), interest value (.11 compared to -.12), and utility value (.14 compared to -.15). Boys also scored higher for math achievement (.06 compared -.07), but the difference here was smaller compared to the motivation variables. Regarding math aspiration, female students scored higher (.01 compared to -.01). The difference however, is negligible.

3.2 Correlations

3.2.1 Total sample.

All correlation tables are presented in the third appendix. Zero-order correlations between the different variables are presented in table C1. Gender was negatively correlated with tracking, indicating slightly less boys in higher tracks. Higher tracks scored higher on all motivation variables, as well as achievement and aspirations. Gender was positively correlated (i.e. male students) with self-efficacy, both value variables, as well as achievement, but not with aspirations. Across the whole sample, self-efficacy and interest value were both positively correlated with achievement but not with math aspirations. Utility value was positively correlated with both achievement and aspirations.

3.2.2 Cross-track correlations.

Correlation matrices for each track are presented in table C2. Self-efficacy correlated positively with both value variables and achievement in all tracks. Concerning aspirations, the correlation with self-efficacy was positive in VMBO-BL (the correlation coefficient was 0.171). In HAVO and VWO however, the correlations were negative (-.109 and -.143), while in VMBO-Kl and GL/TL the correlations were not statistically significant. The correlations between interest and utility value were positive in all tracks. In regards to achievement, the correlations with interest value were not statistically significant in VMBO-BL and -KL, while in VMBO-GL/TL, HAVO, and VWO the correlations were positive (.210, .178, and .280). The correlations between interest value and aspirations were positive in VMBO-BL and -KL (.220 and .142), while they were negative in HAVO and VWO (-.083 and -.258). In VMBO-GL/TL the correlation was not statistically significant. Utility value correlated positively with achievement in most tracks. In VMBO-KL however, the correlation was not significant. The correlations between utility value and aspirations were positive in the vocational tracks. In HAVO however, the correlation was statistically not significant, while in VWO the correlation was negative (-.181). A slightly different trend was visible for the correlation between achievement and aspirations. In the vocational tracks the correlations were statistically insignificant, while in HAVO and VWO the correlations were negative (-.136 and -.217).

3.2.3 Cross-gender correlations.

Correlation matrices of the variables included in the current study are presented in table C3 for males and females separately. For female students, all variables but math aspirations correlated positively and significantly with each other. The only positive correlation for math aspirations was with utility value (.123). The correlation analysis for male students was largely the same as for female students with one notable exception: math aspirations negatively correlated with Math achievement (-.066). A remarkable difference between female and male students was found in the correlation between interest value and math achievement (.225 for female students and .135 for male students).

There were not many differences in terms of the direction of correlations between male and female students, except for the correlation between math achievement and aspirations, which was slightly negative in male students and insignificant for female students. Also for both genders, aspirations were not correlated with either self-efficacy or interest value. However, as shown previously in track specific results, the correlation between self-efficacy and aspirations was positive in VMBO-BL but negative in HAVO and VWO. When combined across tracks, the correlation disappeared. The same pattern was observed for interest value which was not related to aspirations for the total sample nor the gender specific samples. It was however positively correlated with aspirations in VMBO-BL and -KL, but negatively in HAVO and VWO.

3.3 Research question 1

The first research question was: *Is there a main effect of motivation, tracking, and gender on math achievement and aspirations?* To address this question, the predictive effects of gender, tracking, and motivation on math achievement for the total sample were analysed. The model was first estimated with a multilevel random slope model, however, none of the slope residual variance parameters were statistically significant. Thus all subsequent analyses were based on multilevel random intercept models for achievement outcomes (table D1, model 1). Then, the same analysis was run for aspirations outcomes (table D1, model 3).

Male students scored significantly higher than female students (.14). Compared with the baseline VMBO-BL group, all other tracks scored higher on achievement with VWO students having the biggest difference in scores compared to VMBO-BL students (values ranging from 0.36 for VMBO-KL students and 2.04 for VWO students). Concerning the main effects of motivation, self-efficacy had a predictive effect on math achievement (.18). Interest value, although the effect was small, also positively affected achievement (.04). Utility value had no significant effect on math achievement.

For math aspirations (table D1, model 3), there was no significant difference between female and male students. Math aspirations of students in higher tracks were higher, although the differences were only statistically significant in HAVO and VWO (.23 for HAVO students, and .18 for VWO students). Concerning motivation, utility value was a positive predictor (.21). Self-efficacy (-.07) and particularly interest value (-.15) had however a negative effect on math aspirations.

3.4 Research question 2

The second research question was: *Is there a multiplicative effect between self-efficacy and interest value and between self-efficacy and utility value on math achievement and aspiration?* Positive interactions signify that high expectancy and high value lead to high achievement or aspirations, while negative interactions demonstrate that high expectancy and value often lead to low achievement or aspirations. For interactions that are not statistically significant, the effect of the interaction was not present or at least not strong enough to be detected in the current sample size.

The multiplicative effect of self-efficacy and interest value and self-efficacy and utility value on math achievement and aspirations are presented in table D1, model 2 and 4 respectively. In regards to achievement, the interaction between self-efficacy and interest value had no significant effect (See figure G1). The interaction between self-efficacy and utility value (.03) however, positively predicted math achievement: high self-efficacy and utility value led to higher achievement (See figure G2). For aspirations, The effect of the interaction between self-efficacy and interest value was not statistically significant (figure G3). There was however a negative multiplicative effect of self-efficacy and utility value (-.14), which means the joint effect of high self-efficacy and utility value led to lower aspirations among students. As shown in figure G4, the effect of self-efficacy on aspirations decreased as utility value increased, which indicates that for students with higher utility value, self-efficacy actually became less important for aspirations.

3.5 Research question 3

The third research question was: *Is there an interaction effect of tracking and motivation on math achievement and aspirations?* The question examined whether there are differences between tracks regarding the relationship between expectancy, value and outcomes in achievement and aspirations. This research question was answered in four steps. First of all, the main effects of motivation on math achievement and aspirations within the five tracks were analysed. Secondly, the interactions between the tracks and each of the expectancy and value variables were examined. Third of all, the multiplicative effects of expectancy and value within the five tracks were investigated. Lastly, three-way interactions between the tracks and each of the multiplicative effects of expectancy and value were analysed. All analyses were performed while controlling for gender. The results of gender however are not discussed in detail in this section, but rather during the following research question.

3.5.1 The main effects of motivation within the five tracks.

The results of the first analysis are presented in table E1. Self-efficacy positively predicted math achievement in all tracks (values between .12 and .21). The effect was the strongest in HAVO and VWO. In the vocational tracks the effect of self-efficacy was weaker. Interest and utility value had no effect on math achievement, with one exception: interest value positively predicted achievement in VWO (.11).

Self-efficacy had no significant effect on aspirations in four of the tracks (see table E2). In HAVO, self-efficacy (-.15) negatively predicted aspirations. Interest value was a negative predictor for aspirations and this effect became stronger in higher tracks: in VMBO-GL/TL (-.12), HAVO (-.20), and VWO (-.23). Utility value positively predicted aspirations in four tracks (between .25 and .37) but the effect became weaker in higher tracks: in VWO, utility value had no significant effect on aspirations.

3.5.2 The interaction between tracks and motivation.

The second step examined the effect of the interactions between motivation and tracks on math achievement and aspirations, testing differences across tracks observed in the first step. The results are presented in table E5. First, the results for achievement are analyzed (model 1). Since VMBO-BL was the reference group, interactions between other tracks and motivation variables represent the differences in the relationship compared to those in VMBO-BL. Although self-efficacy was a positive predictor for achievement, there were no significant differences in its relationship with achievement between tracks. Concerning interest value, the only statistically significant interaction was between with the VWO track (.12): the predictive effect of interest value was stronger in VWO compared to VMBO-BL. For utility value, there were no significant interactions with tracks.

For predicting math aspirations, there were two significant interactions between motivation and tracks, which were both negative. The interaction between interest value and VWO (-.19) and utility value and VWO (-.39) were negative. This means that interest and utility value had a more negative effect on math aspirations in VWO than in VMBO-BL.

3.5.3 The multiplicative effect of expectancy-value.

In table E3 and E4 the multiplicative effect of expectancy and value on math achievement and aspirations within the tracks are presented. In most cases, the interaction between expectancy and value had no significant effect on achievement. There were however three exceptions. In the VMBO-BL track both interactions had a significant effect. The interaction between self-efficacy and interest value (-.11) negatively affected achievement, whereas the interaction between self-efficacy and utility value (.14) had a positive effect. In the VMBO-KL track the interaction between self-efficacy and interest value (-.08) also negatively affected achievement. Furthermore, there is a pattern visible which points towards a stronger effect of the interaction of self-efficacy and interest value on achievement in higher tracks.

The multiplicative effect of expectancy and value on math aspirations was mostly not statistically significant. However, following a pattern similar to achievement, the interaction between self-efficacy and interest value became more positive in higher tracks. In VWO the interaction between self-efficacy and interest value (.10) positively predicted math aspirations. For the interaction between self-efficacy and utility value, the effect size was strongest in VMBO-KL (-.18) and VMBO-GL/TL (-.15). This indicates that in these tracks for students with high utility values, the predictive effect of self-efficacy further lowers aspirations.

3.5.4 The interaction between tracks and expectancy-value.

Lastly, the three-way interactions between the tracks and the multiplicative effect of expectancy and value are shown in table E6. The three-way interactions serve to confirm the differences in interaction effects across tracks in the previous section. Regarding the effect of the three-way interaction between self-efficacy and interest value and track on achievement, the only significant interaction was in the VMBO-TL/GL track (.13). This was consistent with the previous result that VMBO-BL had the most negative interaction for self-efficacy and interest value whereas VMBO-GL/TL had the most positive interaction (table E3).

Concerning the three-way interaction effects between self-efficacy and utility value and tracks, the only significant effect was detected in the VMBO-GL/TL track. Here however, the effect was negative (-.13), which means that the interaction effect between self-efficacy and utility value on math achievement was significantly lower than in VMBO-BL. This is consistent with the earlier results across tracks that VMBO-BL had the largest interaction, whereas VMBO-GL/TL the least (table E3). For math aspiration no significant interactions between tracks and expectancy-value were found.

3.6 Research question 4

The fourth research question was: *Is there an interaction effect of gender and motivation on math achievement and aspirations?* Exactly like the third research question it was answered in four steps. Firstly, the main effects of motivation on math achievement and aspirations for female and male students were examined. Second of all, the interactions between gender and self-efficacy, interest value, and utility value were analysed. Then, the multiplicative effects of expectancy and value for female and male students were investigated. Lastly, three-way interactions between gender on the one hand and the multiplicative effects of expectancy and value on the other hand were scrutinized. All analyses were performed while controlling for tracking and the interaction of tracking and expectancy-value, although tracking is not the focus of the current research question.

3.6.1 The effects of tracking and motivation for female and male students.

The effects of motivation on math achievement and aspirations for female and male students separately are presented in table F1. First, the results for math achievement are examined. Self-efficacy positively affected achievement for both female (.16) and male students (.20). Interest value also positively predicted achievement for girls (.07), but not for boys. Utility value had no significant effect on achievement for both boys and girls.

The results for aspirations followed a different pattern, as is visible in table F1. Self-efficacy negatively predicted aspirations for boys (-.09), while there was no significant effect for girls. Interest value negatively predicted aspirations for both female (-.18) and male students (-.13), whereas utility value positively predicted aspirations for female (.22) and male students (.20).

3.6.2 The interaction between gender and motivation.

The results of the interactions between gender and motivation are presented in table F2. Self-efficacy (.16) and interest value (.07) positively predicted achievement, while utility value (-.04) negatively predicted achievement. The only significant interaction between gender and motivation was between gender and interest value. The effect of interest value on achievement was stronger for girls (.-06). The other two interactions were not significant, meaning that the effect for male and female students was comparable.

The results regarding math aspiration are also presented in table F2. There was no significant difference in math aspirations for boys and girls. Although self-efficacy was not a predictor of math aspirations, a significant difference between girls and boys was found: the effect for girls was stronger than for boys (-.11). Interest value (-.17) negatively predicted aspirations, while there was no significant difference between female and male students. Utility value positively affected aspirations (.22). There was however no significant difference between both genders.

3.6.3 The multiplicative effects of expectancy-value for female and male students.

The results for math achievement and aspirations are presented in table F3. For both female and male students, there were no significant interactions predicting math achievement. For math aspirations, the results differed somewhat from the model in the first step. Self-efficacy and interest value negatively predicted aspirations for boys and girls. Utility value positively affected aspirations for both genders. The interaction between self-efficacy and interest value had no significant effect for both female and male students. The interaction between self-efficacy and utility value however negatively predicted math aspirations for boys (-.11) and girls (-.19).

3.6.4 The interactions between gender and expectancy-value.

The results for the effects of the interactions between gender and expectancy-value on math achievement and aspirations are presented in table F4. For math achievement, the interaction between self-efficacy and interest value was not significant. There was also no significant difference between female and male students. The interaction between self-efficacy and utility value positively predicted achievement (.14). although there was no significant difference between boys and girls. Concerning math aspirations, both multiplicative effects of expectancy and value were not statistically significant. There were also no significant differences between male and female students for both interactions.

4 Discussion

Using the expectancy-value theory, this study investigated whether the effect of motivation for mathematics achievement and aspirations varied by tracks and gender for 15 year old students in the Netherlands utilizing PISA data from 2012. The present chapter consists of four sections. Firstly, the hypotheses are examined and possible explanations for their verification or falsification are discussed and connected to other studies. Then, the limitations of this study are discussed. Subsequently, practical and policy implications of this study for education in the Netherlands are examined. Lastly, proposals for further research are made.

4.1 The hypotheses

The first research question examined the main effects of tracks, gender, and motivation on math achievement and aspirations. All results are found in Table D1 in the Results section. For achievement, the hypothesis was verified. As expected, higher tracks had increasingly higher scores, although the difference between HAVO and VWO was smaller than the difference between VMBO-GL/TL and HAVO.

Verifying the hypothesis, aspirations were higher in higher tracks. There were no statistically significant differences in aspirations for VMBO-KL and VMBO-GL/TL students compared to VMBO-BL students. In HAVO and VWO aspirations were significantly higher compared to the VMBO-BL track, although the difference between VMBO-BL students and HAVO students was bigger than the difference with VWO students. The results thus confirm earlier studies (Trautwein et al., 2006; Maaz et al., 2008) that students in higher tracks have higher aspirations. In Table D1, a gradual increase of the regression coefficients for the tracks is visible. Higher aspirations in higher tracks may be a result of the fact that in the VMBO tracks all students choose one of the ten available learning profiles, in which mathematics is obligatory in only six profiles, while in HAVO and VWO, math is obligatory for almost all students. Because in VMBO there will be relatively more students without mathematics in their learning profile, students most likely have lower aspirations than HAVO and VWO students. Almost all HAVO and all VWO students are obligated to follow math courses so they are likely to have higher mathematical aspirations.

In regards to gender, the effects on achievement were different than predicted. Despite an abundance of evidence supporting the gender similarities hypothesis (Hyde, 2005), this study found a considerable gap between female and male students in Math achievement, favouring boys. Else-Quest et al. (2010) also found a small but significant difference boys and girls in Math achievement for the TIMSS test in The Netherlands. For the PISA 2003 test there was also a small difference between female and male students which thus still existed in 2012.

Falsifying the hypothesis, no significant difference was found for aspirations between boys and girls. Similar aspirations for both genders is positive as most studies find lower aspirations for female students (Watt et al., 2012; Lauermann et al., 2017). Although women’s share in the STEM field in the Netherlands is still considerably low as mentioned earlier this study, having comparable aspirations is a start in increasing their share.

Consistent with other studies such as Guo et al. (2015), Trautwein et al. (2012), and Nagengast et al. (2011), expectancy (self-concept was used as a measure of expectancy in those studies) was a positive predictor for math achievement. For aspirations, the results were different than predicted and contradicted a multitude of research (Eccles and Wigfield, 2002); there was a negative effect of self-efficacy on aspirations. The negative effect might be explained by the way aspirations were measured. For the first time, PISA utilised the so-called ‘Forced Choice’ format, where students had to choose between Mathematics and either Science or Language in different situations. There are several problems with this format. Firstly, aspirations for either science or languages does not necessarily mean that a student has no mathematical aspirations. However, the ‘Forced Choice’ format implies that it does. Additionally, one of the ‘Forced Choice’ items was ‘I intend to take additional mathematics/language courses after school finishes. If a student has high self-efficacy and even high grades for mathematics, why take extra courses? In the Netherlands it is not common to take additional courses after school for subjects students are good at, but rather for subjects which they find difficult.

Interest value was expected to only predict aspirations. In this sample however, Interest value positively predicted achievement, which was different than Guo et al (2015), who found no significant effect, and Trautwein et al. (2015), where intrinsic value negatively affected achievement. Interest value was however a negative predictor of aspirations, which is different than previous findings by Guo et al. (2015) and Nagengast et al. (2011). A part of the explanation may again be found in the ‘Forced Choice’ format mentioned above.

The results for utility value verified the hypothesis as it was found to be a positive predictor for aspirations, which was similar to the findings of Guo et al. (2015). As expected, no significant effect on achievement was detected, which is similar to findings by Trautwein et al. (2012), but different than Guo et al. (2015), who found a positive effect. Although these results are in line with the hypothesis, it does conflict somewhat with the abovementioned thoughts on the ‘Forced Choice’ format. If the format in fact significantly changes how math aspirations is defined in this study, the effect of utility value would probably also be less pronounced. Clearly, further investigation is needed to conclude whether the ‘Forced Choice’ format changes the purport of math investigation and whether the format is tenable in future PISA research.

Balancing the results of expectancy and value, utility value stood out as the only positive predictor for aspirations for the Dutch 15-year old students, while self-efficacy and interest value were negative predictors. For achievement it was the opposite: self-efficacy and interest value were positive predictors, while utility value had no effect. It is difficult to assess to which extent the ‘Forced Choice’ format is responsible for these results, because its mechanisms are unclear and difficult to explain, and further research on the subject is needed, which is elaborated on further on in the discussion section.

The second research question investigated whether there was a significant multiplicative effect of expectancy-value on math achievement and aspirations. All results are found in Table D1 (models 2 and 4) in the Results section. The interaction effects of self-efficacy and interest value on achievement and aspirations were not statistically significant, which was the same finding as Guo et al. (2015) in the Hong Kong context. It was however different than for example and Nagengast et al. (2011), where a positive effect on math aspirations was found, and Trautwein et al. (2012) who detected a positive effect on Math achievement in the German context.

Concerning the interaction between self-efficacy and utility value, a positive effect on achievement was found; the effect of utility value on achievement amplified as self-efficacy increased (Figure G2), which was in line with other studies such as Trautwein et al. (2012) but different than for example Guo et al. (2015), who found a negative effect on achievement with Hong Kong students. It is clear that self-efficacy was the strongest predictor of achievement. Remarkably, for aspirations, the coefficient of the interaction between self-efficacy and utility value was negative. Here, as utility value increased, the effect of self-efficacy weakened, as visualised in figure G4. For aspirations, utility value was the strongest predictor. Although a positive effect was expected, these findings are similar to those of Guo et al. (2015), where the effect of self-concept on math aspirations also was stronger at lower levels of and utility value. This may in part be due to the previously discussed issue with the measurement of aspirations where higher aspiration measured by taking further courses could actually be an indication of low self-efficacy. Although an expected positive interaction had a negative effect in this specific context, it does represent a multiplicative effect of expectancy and value.

The third research question looked into whether interaction effects of tracking and motivation on math achievement and aspirations were present in the sample and examined whether differences of the EVT predictions were similar across the various tracks. In order to answer the question, four steps in the analysis were taken. Firstly, the main effects of motivation on math achievement and aspirations within the five tracks were analysed and during the second step the interaction effects of tracking and motivation on achievement and aspirations were analysed in order to detect statistically significant differences between tracks. In all tracks, self-efficacy positively predicted achievement (table E1) and no significant differences were found between tracks in the interaction model (table E5). Concerning aspirations, self-efficacy only had an effect in HAVO, where a negative effect was detected (table E2). The results of the second step (using VMBO-BL as a baseline; table E5) however, showed that there were no significant differences in the effect of self-efficacy on aspirations.

Interest value only had a positive effect on achievement in VWO (table E1). The interaction model confirmed that only the effect of interest value in VWO significantly differed from VMBO-BL (table E5). This contradicts Frenzels (2010) finding that students in lower tracks have more favourable interest trajectories (i.e. higher levels of interest concerning mathematics). In regards to aspirations, interest value was a negative predictor in all tracks, but only statistically significant in the three highest tracks (table E2). The interaction model revealed that only the effect of interest value in VWO significantly differed from the effect in VMBO-BL (table E5, model 2). The negative effect of interest value on math aspirations is quite remarkable and warrants further research into why math interest doesn't lead to higher math aspirations.

Although no effect of utility value on achievement in any of the tracks was expected, a positive effect was found in VMBO-BL. In VMBO-KL and VMBO-GL/TL the effect was also positive but not statistically significant (table E1). The interaction model confirmed this finding as the effect of utility value was weaker in all other tracks compared with VMBO-BL. Particularly in VWO where the weaker effect was statistically significant (table E5, model 1). In line with an abundance of evidence (Wigfield and Eccles, 2002), utility value had a positive effect on aspirations in all tracks. Remarkably, the predictive effect of utility value on aspirations decreased in higher tracks (Table E2). The interaction model however demonstrated that only the results in VWO were significantly weaker than in VMBO-BL (Table E5, model 2). Given the fact that mathematics is obligatory for all students in VWO, and in three of four profiles in HAVO, it is perhaps less surprising that the effect of utility value is stronger in vocational tracks. In the VMBO tracks students only follow math courses if their learning profile requires it. These students are naturally more likely to pursue math invested careers as the learning profile prepares students for their further vocational education. The almost absent effect of utility value on achievement is in line with earlier findings by for example Trautwein et al. (2012), but different than Guo et al. (2015), who found a positive effect of utility value on both achievement and aspirations.

Then, the multiplicative effects of expectancy-value on math achievement and aspirations within and between tracks were analysed. Regarding achievement, only three statistically significant interactions were found. Within the VMBO-BL and -KL track, the interaction of self-efficacy and interest value was a negative predictor (table E3). The analysis between tracks (table E6) demonstrated that the effect of the interaction was higher in all tracks compared to VMBO-BL, but only in VMBO-GL/TL the effect was significantly higher. This was also the only track where the interaction had a positive sign, but however not statistically significant. The interaction between self-efficacy and utility value had a positive effect on achievement in VMBO-BL (table E3); the only result confirming the hypothesis. The results for the analysis between tracks (table E6) showed that the interaction in VMBO-GL/TL was significantly weaker than in VMBO-BL. Although the effect in the other tracks was also weaker than in VMBO-BL, the differences were not statistically significant.

Concerning aspirations, the interaction between self-efficacy and interest value was a positive predictor in VMBO-, -KL, HAVO, and VWO, but only statistically significant in the latter (table E4). In all other tracks the interaction was not statistically significant. In the last step of the analysis no significant differences between tracks were found (table E6). The results for the interaction between self-efficacy and utility value were also different than predicted. No positive effects were found. In VMBO-KL and -GL/TL a negative effect on aspirations was found (table E4). Between tracks however, no significant differences were found (table E6). Concluding, the multiplicative effect of expectancy-value does not seem to vary by tracking.

The last research question concerned the effects of gender and motivation on math achievement and aspirations. Their effects were again analysed in four steps. First of all, the main effects of tracking and motivation on math achievement and aspirations were explored for female and male students separately. Subsequently, the interactions between gender and motivation were examined. In regards to achievement, the results demonstrated that tracking was a significant predictor for both genders, although the effect for female students was somewhat stronger than for their male peers (table F1).

Self-efficacy had a positive effect on achievement for boys and girls (table F1 for all motivational behaviours). The analysis of the interactions however, showed no significant difference in the effect of self-efficacy (See table F2 for all interaction analyses). Guo et al. (2015) found that male students had higher levels of self-concept and because of its positive effect on achievement, it led to higher achievement. When female and male students had similar levels of self-concept however, girls achieved better. For girls, interest value also positively predicted achievement, whereas for boys no significant effect was found. The interaction between gender and interest value confirmed a significant difference between female and male students; interest value had a stronger effect on achievement for girls in this sample. Regarding interest value, Guo et al. (2015) found no significant differences between both genders. The same applies to utility value. In this study, utility value also had no significant effect on achievement for boys or girls and no significant difference between genders was found in the interaction analysis.

Concerning aspirations (see again Table F1 for all motivational behaviours), there were no statistically significant differences between tracks for male students. For female students however, aspirations increased in higher tracks, as predicted. This could be a result of the fact that more girls enrolled in HAVO and VWO choose learning profiles in which math is obligatory compared to girls in VMBO: in the academic year 2016/2017 almost four times as many girls (5706 girls and 1552 boys) chose the profile in HAVO where mathematics is not obligatory.[7] In the same year, in VMBO approximately 1293 girls (4% of 32.336 girls in total) were enrolled in technical learning profiles, compared to more than 12.000 boys (32% of 37.624 boys in total).[8] Different than the hypothesis, self-efficacy had a negative effect for male students, whereas the effect for female students was not statistically significant. The interaction analysis (See table F2 for all interaction analyses) conformed these findings as it showed a significant difference between boys and girls favouring the latter. Interest value was expected to positively affect aspirations for boys and girls. The results however showed a negative effect for both female and male students. The second analysis showed no significant difference in the effect of interest value on aspirations between females and males. The results for utility value verified the hypothesis, as it positively predicted aspirations for both genders. The analysis demonstrated that the

[7] See: https://www.vhto.nl/cijfers-onderzoek/cijfers/cijfers-havovwo/

[8] https://www.vhto.nl/cijfers-onderzoek/cijfers/cijfers-vmbo/

effect of utility value was evenly strong for both genders. These findings are quite different than those by Guo et al. (2015), who first of all found that girls had higher aspirations, secondly that expectancy and utility value were stronger predictors of aspirations for boys, and lastly that interest value was an evenly strong predictor of aspirations for both genders.

During the third and fourth step of the analysis the multiplicative effect of expectancy-value within genders (table F3) and then the interactions of gender and expectancy-value were examined (table F4). Firstly, the interaction between self-efficacy and interest value was analysed. Despite the hypothesis, no effects were found for achievement or aspirations, and no differences were found between female and male students in the fourth analysis (tables F3 and F4). However, no statistically significant difference of the multiplicative effect was found between boys and girls. Clearly contradicting the hypothesis, the interaction between self-efficacy and utility value negatively predicted aspirations for both girls and boys (table F3). No significant difference was however found between both genders was in the fourth analysis (table F4). In other words, the multiplicative effect of expectancy-value did not vary by gender.

4.2 Limitations of the present investigation

The current study had several limitations. One limitation is that the present investigation only considered tracks alone. It is also important to explore the effect of the variation in the composition of tracking in secondary schools in the Netherlands, because currently there are no clear rules or guidelines for structuring the various tracks in different schools. It is possible that a school which offers multiple tracks, situates the different tracks in different buildings, different sections of buildings, or mix all or some of the tracks in one building or even in a single classroom. All these options are likely to have an effect on student motivation and educational outcomes. Which effect however is unclear. For example, it may be useful to compare the methods of organising tracking including comparing against schools from what in the Netherlands are called ‘broad school communities’, schools that offer all tracks.

A second aspect which may have influenced the results is the fact that not all students in the sample followed a mathematics course. In the vocational tracks, relatively fewer students followed a math course compared to HAVO and VWO. In HAVO, students from three of four learning profiles followed a math course, while in VWO every student followed a math course. As discussed in the previous section more girls tend to choose the profile without math (see footnote 1 and 2). Ideally, the sample would only consist of students that follow math courses or using whether or not following a math course as a controlled variable. The PISA 2012 Technical Report does not answer the question whether this indeed is the case. What further complicates this matter is that VMBO students choose their profile in the third year, while HAVO and VWO students do the same in the beginning of their fourth year. The sample is made out of 15 year old students who are either in their third or fourth year.

This implies that all VMBO students in the sample had chosen their learning profile. In this dataset 873 (41,5%) HAVO and VWO students were in their third year, or ninth grade, and had not chosen a profile yet, and thus followed a math course. 1228 (58,5%) HAVO and VWO students were in their fourth year and hence had chosen a learning profile at the time the PISA test was taken.

A third problem in this study is the so called ‘Forced choice’ format, which PISA used to measure mathematic aspirations, and in particular its relation to the Dutch context. The ‘Forced Choice’ format was introduced by PISA in the 2012 test. Some of the forced choices were ambiguous in the Dutch context, particularly because it was not made clear to students that the ‘Forced Choice’ questions were used to measure mathematic aspirations. For instance, when choosing between taking additional math or language classes after school finishes (ST48Q01), a Dutch student was perhaps more likely to choose the subject which he found more difficult instead of the subject for which he had career aspirations. In the Netherlands, taking extra classes during or outside of school is almost exclusively done in order to improve results which are not yet good enough. ST48Q03 is also ambiguous because students are forced to choose whether they are willing to study harder for mathematics or language classes. Because it was not made clear to students that the items were about aspirations, they might again have chosen for the subject they found more difficult. Choosing these two answers may have contradicted the answers of the two other items (ST48Q02 and ST48Q05) which were clearly about math aspirations. The fourth Item (ST48Q04) forced students to choose between taking as many science or math classing during his or her education. This item is as much about interest as it is about aspirations. These ambiguities might explain the non-existent or low correlation with the other variables and thus also why many hypotheses concerning the multiplicative effect of Expectancy-value were falsified. The items measuring aspiration were for instance quite different from Nagengast et al. (2012) and Guo et al. (2015). The former assessed extra-curricular activities on the one hand and students’ intentions of studying science after high school and whether or not aspiring a science invested career. In Guo et al. (2015) the only item assessing aspirations asked students how far in school they expected to go.

A further limitation of this study is that it is plausible that there were unmeasured or ignored variables which have had an effect on the outcome variables. An example of an unmeasured variable is for instance students’ prior academic achievement. Examples of ignored variables, which thus were available but ignored for practical matters, are educational aspirations and academic ability of parents (which both significantly affect students’ achievement and aspirations), the social background of students, anxiety for mathematics, and attributions to failure, all of which impact student achievement and aspirations (Dockx et al., 2018).

Also, when investigating whether motivation varies by tracking, comparing levels of motivation, achievement, and aspirations of students before and after tracking would provide a lot of insight in the effects of tracking. Unfortunately, the data collected and offered by PISA did not provide the opportunity to do this.

Lastly, the data is cross-sectional, which implies limited validity in terms of causal inference. Future studies based on longitudinal data, or randomised intervention are required to draw conclusions in stronger causal terms, Nevertheless, the data is population based, thus representing the target population well with minimal sampling bias. Furthermore the sample size is large and therefore provides strong power for the statistical analysis and results presented in the present study.

4.3 Practical and policy implications

The main goal of this study was to investigate how student motivation varied by tracking and gender and to what extent motivation affected achievement and aspirations in a mathematical context. The mathematical context was particularly relevant due to the growing shortages of workforce and the lack of women in the STEM fields in the Netherlands.

The most important conclusion of this study is that student motivation was consistently predictive of achievement and aspirations. Regarding achievement, there was not much difference across the tracks. Self-efficacy was predictive across tracks in a comparable way. Only interest value was more predictive in VWO than in other tracks. Utility value on the other hand had no effect in any of the tracks. This implies that it might be effective to target self-efficacy to improve student achievement across all tracks. In VWO, increase interest in math might improve achievement. This may be due to the academic-oriented nature of the VWO track in which intrinsic motivation (i.e. interest) is more likely associated with achievement in comparison to other tracks, where vocational education is more the focus of the educational curriculum.

Concerning aspirations, self-efficacy and interest value seemed to be more predictive in higher tracks, while the effect of utility value seemed to decrease. This suggests that it may be especially effective to target utility value for enhancing students’ math career aspiration in vocational tracks, which is in line with the nature of the practical-orientated study programme for this students. Although the negative effects of self-efficacy and interest in higher tracks are surprising, it might indicate that in the more academically focused tracks, students who are interested or confident in maths are not those focused on a career or further training in maths. Nevertheless, future studies are required to clarify these findings in order draw policy implications.

Regarding the multiplicative effects of expectancy-value, there were also differences across tracks. Only one of four interactions tested in the second research question was positive (table 5), namely the effect of self-efficacy and utility value on math achievement. Its effect was more pronounced in lower tracks. The interaction effect of self-efficacy and interest value on achievement was not statically significant in higher tracks, but was stronger (i.e. more negative) in lower tracks. Although the EVT predictions presume a synergetic relationship between expectancy and value, the findings in the current study are more varied than hypothesised. The negative interaction of self-efficacy and interest value predicting achievement in lower tracks may imply that interest does play a role in the expectancy-achievement prediction, but this relationship becomes weaker as student interest increases. Perhaps due to the practical-oriented nature of the educational programme in vocational tracks, although the mechanisms of this relationship needs further investigation. More consistent with the nature of tracking and EVT predictions is the positive interaction between expectancy and value for the vocational track VMBO-BL. This suggests that it may be particularly effective to target both expectancy and utility value in order to improve math achievement for VMBO-BL students.

Concerning aspirations, the interaction effect of self-efficacy and interest value and of self-efficacy and utility value, seemed to be more positive in the higher tracks. The positive expectancy by interest interaction amongst VWO students is in line with EVT and tracking predictions where intrinsic motivation measured by interest may indeed form a synergetic relationship with self-efficacy for students pursuing future math career and training. However the negative interaction with utility value in lower track students is counter intuitive. Further investigation is required in order to better understand this finding and its implications.

In regards to tracking, the results demonstrated that tracking has a major effect on achievement and aspirations. Both outcomes were higher in higher tracks. Lower aspirations in lower tracks mean that these students may indeed be less likely to pursue math related careers and further education, which may hinder the supply of labour market from these tracks. In order to increase math aspirations amongst all students and reduce the gap in math achievement, a potential solution is making mathematics obligatory in all learning profiles in all tracks. Furthermore, in this study self-efficacy and interest value were negative predictors of aspirations. It’s difficult to explain why confident and interested students do not pursue careers in the STEM fields. Although the problematic ‘Forced choice’ format partly explains the results, there is also an important role for mathematics teachers in stimulating these students to pursue math invested careers, particularly because interest had such a negative effect on math aspirations in this dataset. The results suggest that there are many students interested in mathematics who pursue careers in other fields.

In regards to gender, although this study showed that there is still a gap in the Netherlands concerning achievement in mathematics, the effects of motivation on achievement and aspirations were quite similar across genders. Detected differences were the effect of self-efficacy on aspirations which had no effect for female students but was negative for male students. Also, interest value positively predicted achievement for girls but had no effect for boys. In both cases, the difference between both genders was statistically significant. While self-efficacy was equally effective for both boy and girls, it may be especially effective to target interest for girls in order to improve achievement. The interactions between expectancy and values did not differ across genders. Although the negative effect of self-efficacy (only for boys) and interest on aspirations was somewhat surprising, the positive effect of utility value implies that it may serve as an effective intervention target in order to increase math aspirations, for both boys and girls.

4.4 Proposals for further research

Given the results of this study, there are various proposals for further research. First of all, it would be informative to reduce some of the limitations mentioned earlier in the discussion, such as the composition of secondary schools in the Netherlands. Does motivation vary across different school compositions? Another problem was whether students followed math courses or not. Although it is likely that most of the students did follow a math course, knowing they did could refine the results. Additionally, including data on motivational behaviour before students are tracked would prove insightful. This data could for instance be gathered together with the exit test during the last year of primary school. Collecting data before tracking is useful for other insights as well. This study demonstrated that a gender gap in mathematics, for motivation and achievement, still exists in the Netherlands. It could be informative to explore how and where the gender gap originates. Does this occur before or after tracking? If before tracking, at what age does the gap come into being?

Moreover, comparing different countries where early tracking based on academic ability is common, in order to explore whether tracking affects student motivation similarly in different contexts would be very interesting. Moreover, comparing the Dutch context with Western countries where students are tracked at a later age, such as the Scandinavian countries, could shed light on the development of motivational behaviour of students and the effect of tracking. On the one hand comparing Western countries with each other could offer important insights, but also cross-cultural comparisons could prove very insightful.

Thirdly, it would be insightful to expand the current research by adding more variables in order to refine the results. Adding the socioeconomic status of students as a variable would be very interesting as tracking is widely accepted as an important factor increasing social inequalities. (Van de Werfhorst & Mijs, 2010). Also adding math anxiety could clarify the results, as anxiety is more common amongst female students (Stoet et al., 2016). Therefore, math anxiety might have a significant effect on motivation and outcomes of girls, which has not been taken into account in this study.

Furthermore, it would be interesting to explore whether self-efficacy is a negative predictor of aspirations only in this sample or whether it is something common in the Netherlands. Why do feelings of competency lead to lower career aspirations, particularly in higher tracks? The same applies to interest value. Is its negative effect on math aspirations confined to this sample or is it common in the

Netherlands? If the negative relation between self-efficacy and interest value and aspirations is common in the Dutch context, further exploring it, for example by researching their origins and effects before students are tracked, could offer important insights for teachers.

Concluding, nowadays there is a rich body of literature about the effects of tracking. At the same time with every question answered or explored, new questions arise to the surface. Therefore continuing research on the effects of tracking is indispensable.

5 Literature

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological review*, *64*(6), 359-372.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, United States: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, United States: Freeman.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child development*, *72*(1), 187-206.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173-1182.

Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter?. *Journal of educational psychology*, *104*(3), 682-699.

Bong, M. (2001). Between-and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task value, and achievement goals. *Journal of educational psychology*, *93*(1), 23-34.

Branden, N. (1995). *The six pillars of self-esteem*. New York, United States: Bantam Books.

Brookover, W. B., & Lezotte, L. W. (1979). *Changes in school characteristics coincident with changes in student achievement*. East Lansing, United States: The Institute for Research on Teaching, Michigan State University.

Calsyn, R., & Kenny, D. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology*, *69*(2), 136-145.

Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction*, *18*(4), 321-336.

Chouinard, R., & Roy, N. (2008). Changes in high-school students' competence beliefs, utility value and achievement goals in mathematics. *British journal of educational psychology*, *78*(1), 31-50.

DeBacker, T. K., & Nelson, R. M. (2000). Motivation to learn science: Differences related to gender, class type, and ability. *The Journal of Educational Research*, *93*(4), 245-254.

De Corte, E., & Op 't Eynde, P. (2003). *When girls value mathematics as highly as boys: an analysis of junior-high students' mathematics-related beliefs*. Paper presented at 2003 Annual Meeting of the American Educational Research Association, Chicago, United States of America. Retrieved from https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1938686&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US&fromSitemap=1

de la Fuente, J., Sander, P., & Putwain, D. (2013). Relationship between undergraduate student confidence, approach to learning and academic performance: The role of gender. *Revista de Psicodidáctica*, *18*(2), 375-392.

Denissen, J. J., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child development*, *78*(2), 430-447.

Dockx, J., De Fraine, B., & Vandecandelaere, M. (2018). Does the Track Matter? A Comparison of Students' Achievement in Different Tracks. *Journal of Educational Psychology*. Obtained from http://www.dx.doi.org/10.1037/edu0000305 on January 4th

Dronkers, J. (2013). Kleinere rol Cito bevoordeelt 'ons soort mensen'. *NRC Handelsblad*. Attained on 16 May 2018, http://apps.eui.eu/Personal/Dronkers/Dutch/NRC2013.pdf

Dupriez, V., Monseur, C., Van Campenhoudt, M., & Lafontaine, D. (2012). Social Inequalities of Post-secondary Educational Aspirations: influence of social background, school composition and institutional context. *European Educational Research Journal*, *11*(4), 504-519.

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*(2), 382-393.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, *53*(1), 109-132.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80, 501-519.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological bulletin*, *136*(1), 103-127.

Field, A. P. (2005) *Is the meta-analysis of correlation coefficients accurate when population correlations vary?* Psychological Methods, 10 (4). pp. 444-467.

Forgasz, H. J., & Leder, G. C. (2017). Persistent gender inequities in mathematics achievement and expectations in Australia, Canada and the UK. *Mathematics Education Research Journal*, *29*(3), 261-282.

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507-537.

Gallagher, A. M., & Kaufman, J. C. (Eds.). (2004). *Gender differences in mathematics: An integrative psychological approach*. Cambridge University Press.

Green, J., Nelson, G., Martin, A. J., & Marsh, H. (2006). The Causal Ordering of Self-Concept and Academic Motivation and Its Effect on Academic Achievement. *International Education Journal*, *7*(4), 534-546.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: developmental perspectives on their causal ordering. *Journal of educational psychology*, *95*(1), 124-136.

Guill, K., Lüdtke, O., & Köller, O. (2017). Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study. *Learning and instruction*, *47*, 43-52.

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences*, *37*, 161-168.

Hackett, G., & Betz, N. E. (1981). A self-efficacy approach to the career development of women. *Journal of vocational behavior, 18*(3), 326-339.

Hanushek, E. A., & Wößmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal*, *116*(510), 63-76.

Hyde, J. S. (2005). The gender similarities hypothesis. *American psychologist*, *60*(6), 581-592.

Inspectie van het Onderwijs. (2016). *De staat van het onderwijs: Hoofdlijnen uit het onderwijsverslag 2016/2017*. Utrecht.

Inspectie van het Onderwijs. (2017). *De staat van het onderwijs: Hoofdlijnen uit het onderwijsverslag 2015/2016*. Utrecht.

Ireson, J., & Hallam, S. (1999). Raising standards: is ability grouping the answer?. *Oxford review of education*, *25*(3), 343-358.

Ireson, J., Hallam, S., & Hurley, C. (2005). What are the effects of ability grouping on GCSE attainment?. *British educational research journal*, *31*(4), 443-458.

Kalaycioglu, D. B. (2015). The Influence of Socioeconomic Status, Self-Efficacy, and Anxiety on Mathematics Achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA. *Educational Sciences: Theory and Practice*, *15*(5), 1391-1401.

Korthals, R. A. (2012). Selection and tracking in secondary education: A cross country analysis of student performance and educational opportunities.

Korthals, R. A., & Dronkers, J. (2016). Selection on performance and tracking. *Applied Economics*, *48*(30), 2836-2851.

Lauermann, F., Tsai, Y. M., & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy–value theory of achievement-related behaviors. *Developmental psychology*, *53*(8), 1540-1559.

Lavrijsen, J., Nicaise, I., & Wouters, T. (2013). Vroege tracking, kwaliteit en rechtvaardigheid. Wat het wetenschappelijk onderzoek ons leert over de hervorming van het secundair onderwijs.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*(2), 184-196.

Lucas, S. R. (1999). *Tracking inequality. Stratification and mobility in American high schools.* New York: Teachers College Press.

Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, *2*(2), 99-106.

Marsh, H. W. (1988). The Failure of Academically Selective High Schools To Deliver Academic Benefits: The Importance of Academic Self-Concept and Educational Aspirations.

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of educational psychology, 80*(3), 366-380.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646-656.

Marsh, H. W., & Craven, R. (1997). Academic self-concept: Beyond the dustbowl. *Handbook of classroom assessment: Learning, achievement, and adjustment*, 131-198.

Marsh, H. W., & Hau, K. T. (2003). Big-Fish--Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American psychologist*, *58*(5), 364-376.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397-416.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*(2), 133-163.

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin*, *34*(4), 542-552.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K. T., & Trautwein, U. (2011). Who took the "×" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science*, *22*(8), 1058-1066.

Nagengast, B., & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, *104*(4), 1033-1053.

Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational research and Evaluation*, *12*(4), 323-345.

Nagy, G., Watt, H. M., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The Development of Students' Mathematics Self-Concept in Relation to Gender: Different Countries, Different Trajectories?. *Journal of Research on Adolescence*, *20*(2), 482-506.

OECD (2013), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, OECD Publishing.

OECD (2014), *PISA 2012 Technical Report*. OECD Publishing.

OECD (2014), *Education at a Glance. OECD Indicators*. OECD Publishing.

Pajares, F., & Schunk, D. (2001). The development of academic self-efficacy. *Development of achievement motivation. United States, 7*.

Pisa, O. E. C. D. (2007). Science competencies for tomorrow's world. *Executive Summary. Executive Summary*.

Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2011). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education*, *95*(3), 458-476.

Rubin, D.B. (1987), *Multiple Imputation for Nonresponse in Surveys*, John Wiley & Sons, New York.

Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), Self-efficacy, adaptation, and adjustment: Theory, research, and application (pp. 281-303). New York: Plenum Press.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of educational research*, *46*(3), 407-441.

Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *Policy Insights from the Behavioral and Brain Sciences, 2(1),* 4–12.

Svinicki, M. D. (2016). Motivation: An Updated Analysis. *IDEA Paper, 59,* 1-9.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, *98*(4), 788-806.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of educational psychology*, *104*(3), 763-777.

Tzohar-Rozen, M. & Kramarski, B. (2014) Metacognition, Motivation, and Emotions: Contribution of Self-Regulated Learning to Solving Mathematical Problems. *Global Education Review, 1(4),* 76-95.

Van de Werfhorst, H. & Mijs, J.B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology (36)*, 407-428.

Van Houtte, M. (2017). Gender differences in context: the impact of track position on study involvement in flemish secondary education. *Sociology of Education*, *90*(4), 275-295.

Van Houtte, M., & Stevens, P. A. (2009). Study involvement of academic and vocational students: Does between-school tracking sharpen the difference?. *American Educational Research Journal*, *46*(4), 943-973.

Van Houtte, M., Demanet, J., & Stevens, P. A. (2012). Self-esteem of academic and vocational students: Does within-school tracking sharpen the difference?. *Acta Sociologica*, *55*(1), 73-89.

Wang, M. T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental psychology*, *48*(6), 1643-1657.

Wang, M. T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, *29*(1), 119-140.

Wang, M. T., & Eccles, J. S. (2012). Adolescent behavioral, emotional, and cognitive engagement trajectories in school and their differential relations to educational success. *Journal of Research on Adolescence*, *22*(1), 31-39.

Watt, H. M. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educational Research and Evaluation*, *12*(4), 305-322.

Watt, H. M., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental psychology*, *48*(6), 1594-1611.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological review*, *92*(4), 548-573.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of educational psychology*, *89*(3), 451-469.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary educational psychology*, *25*(1), 68-81.

Zimmerman,B.J. (2002) Becoming an self-regulated learner: An overview. *Theory Into Practice*, 41, 64-70.

Appendix A Survey items of the constructs

Interest value

1. I enjoy reading about mathematics
2. I look forward to my mathematics lessons
3. I do mathematics because I enjoy it
4. I am interested in the things I learn in mathematics

Utility value

1. Making an effort in mathematics is worth it because it will help me in the work that I want to do later on
2. Learning mathematics is worthwhile for me because it will improve my career <prospects, chances>
3. Mathematics is an important subject for me because I need it for what I want to study later on
4. I will learn many things in mathematics that will help me get a job

Self-efficacy

1. Using a <train timetable> to work out how long it would take to get from one place to another
2. Calculating how much cheaper a TV would be after a 30% discount
3. Calculating how many square metres of tiles you need to cover a floor
4. Understanding graphs presented in newspapers
5. Solving an equation like $3x+5= 17$
6. Finding the actual distance between two places on a map with a 1:10 000 scale
7. Solving an equation like $2(x+3) = (x + 3) (x - 3)$
8. Calculating the petrol consumption rate of a car

Mathematic aspirations ('forced choice' format)

1.

a) I intend to take additional mathematics courses after school finishes

b) I intend to take additional <test language> courses after school finishes

2.

a) I plan on majoring in a subject in <college> that requires mathematics skills

b) I plan on majoring in a subject in <college> that requires science skills

3.

a) I am willing to study harder in my mathematics classes than is required

b) I am willing to study harder in my <test language> classes than is required

4.

a) I plan on <taking> as many mathematics classes as I can during my education

b) I plan on <taking> as many science classes as I can during my education

5.

a) I am planning on pursuing a career that involves a lot of mathematics

b) I am planning on pursuing a career that involves a lot of science

Appendix B Descriptive statistics

Table B1

Descriptive statistics for the total sample, for female and male students separately, and for each of the five tracks separately

Descriptives	Total	Girls	Boys	VMBO-BL	VMBO-KL	VMBO-GL/TL	HAVO	VWO
N	4224	2058	2166	396	555	1172	1088	1013
Valid N	2701	1316	1385	232	332	752	726	659
1 Self-efficacy	-0.18	-0.22	0.21	-0.61	-0.46	-0.11	0.08	0.52
2 Interest value	-0.34	-0.12	0.11	-0.08	-0.20	-0.02	0.06	0.11
3 Utility value	-0.38	-0.15	0.14	-0.17	-0.17	-0.03	0.04	0.15
4 Math achievement	525.65	-0.07	0.06	-1.36	-0.94	-0.37	0.42	1.02
5 Math aspirations	0.10	0.01	-0.01	-0.15	-0.01	-0.07	0.09	0.04

Note. Results of the descriptive statistics were obtained from standardized values of each variable excluding ‘gender’. The scores for the Total sample are on the original scale given by PISA.

Appendix C Correlations

Table C1

Correlations total sample

	Gender	Tracks	Self-efficacy	Interest value	Utility value	Math achievement	Math aspirations
1 Gender	x	-0,056***	0,216***	0,115***	0,143***	0,066***	-0,010
2 Tracks		x	0,355***	0,086***	0,108***	0,784***	0,056**
3 Self-efficacy			x	0,389***	0,397***	0,465***	-0,006
4 Interest value				x	0,635***	0,185***	-0,027
5 Utility value					x	0,186***	0,103***
6 Math achievement						x	-0,022
7 Math aspirations							x

*p < 0.05, **p<0.01, ***p<0,001

Note. Results of the correlation analysis were obtained from standardized values of each variable excluding 'Gender' and 'Tracks'

Table C2

Correlations within the different tracks

	Self-efficacy	Interest value	Utility value	Math achievement	Math aspirations
VMBO-BL					
Self-efficacy	x	**0,453*** **	**0,420*** **	**0,275*** **	**0,171** **
Interest value		x	**0,622*** **	0,071	**0,220** **
Utility value			x	**0,125* **	**0,397*** **
Math achievement				x	0,062
Math aspirations					x
VMBO-KL					
Self-efficacy	x	**0,377*** **	**0,428*** **	**0,268*** **	0,074
Interest value		x	**0,631*** **	0,090	**0,142** **
Utility value			x	0,085	**0,269*** **
Math achievement				x	-0,089
Math aspirations					x
VMBO-GL/TL					
Self-efficacy	x	**0,394*** **	**0,411*** **	**0,322*** **	0,053
Interest value		x	**0,623*** **	**0,210*** **	0,058
Utility value			x	**0,192*** **	**0,203*** **
Math achievement				x	-0,042
Math aspirations					x
HAVO					
Self-efficacy	x	**0,388*** **	**0,375*** **	**0,357*** **	**-0,109** **
Interest value		x	**0,657*** **	**0,178*** **	**-0,083* **
Utility value			x	**0,167*** **	0,065
Math achievement				x	**-0,136*** **
Math aspirations					x
VWO					
Self-efficacy	x	**0,354*** **	**0,334*** **	**0,359*** **	**-0,143*** **
Interest value		x	**0,619*** **	**0,280*** **	**-0,258*** **
Utility value			x	**0,181*** **	**-0,181*** **
Math achievement				x	**-0,217*** **
Math aspirations					x

Note. *p < 0.05, **p<0.01, ***p<0.001

Table C3

Correlations for male and female students

	Self-efficacy	Interest value	Utility value	Math achievement	Math aspirations
Female students					
Self-efficacy	x	**0,428*** **	**0,400*** **	**0,473*** **	0,039
Interest value		x	**0,605*** **	**0,225*** **	-0,034
Utility value			x	**0,187*** **	**0,123*** **
Math achievement				x	0,027
Math aspirations					x
Male students					
Self-efficacy	x	**0,328*** **	**0,361*** **	**0,454*** **	-0,041
Interest value		x	**0,652*** **	**0,135*** **	-0,018
Utility value			x	**0,170*** **	**0,089*** **
Math achievement				x	**-0,066***
Math aspirations					x

Note. *p < 0.05, **p<0.01, ***p<0.001

Appendix D Research question 1 and 2

Table D1

The main effects of motivation, gender, and tracking and the multiplicative effects of expectancy-value on math achievement and aspirations. Results from multilevel regression analysis predicting PISA math achievement and aspirations

	Math achievement				Math aspirations			
Predictor	Model 1		Model 2		Model 3		Model 4	
Student Level	β	p	β	p	B	p	β	p
Intercept	**-1,24**	**0**	**-1,25**	0	-0,12	0,09	-0,08	0,28
Female students	Baseline		Baseline		Baseline		Baseline	
Male Students	**0,14**	**0**	**0,14**	0	-0,03	0,46	-0,02	0,62
VMBO-BL	Baseline		Baseline		Baseline		Baseline	
VMBO-KL	**0,36**	**0**	**0,36**	0	0,12	0,17	0,12	0,17
VMBO-GL/TL	**0,81**	**0**	**0,81**	0	0,09	0,27	0,08	0,28
HAVO	**1,58**	**0**	**1,58**	0	**0,23**	**0,004**	**0,22**	0,01
VWO	**2,04**	**0**	**2,03**	0	**0,18**	**0,03**	**0,19**	0,02
Self-efficacy	**0,18**	**0**	**0,19**	0	**-0,05**	**0,03**	**-0,07**	0,002
Value-interest	**0,04**	**0,001**	**0,05**	0,001	**-0,14**	**0**	**-0,15**	0
Value-utility	-0,01	0,29	-0,02	0,25	**0,21**	**0**	**0,21**	0
Self-efficacy*interest			-0,01	0,43			0,02	0,43
Self-efficacy*utility			**0,03**	0,03			**-0,14**	0

Residual variance	Model 1		Model 2		Model 3		Model 4	
	β	p	β	p	β	p	β	p
Level 2 – slope								
Level 2 – intercept	**0,07**	**0**	**0,07**	**0**	**0,02**	**0,03**	**0,02**	**0,02**
Level 1	**0,26**	**0**	**0,26**	**0**	**0,95**	**0**	**0,92**	**0**

*Note.*VMBO-BL is used as reference group. Results are based on standardized regression coefficients for all variables except for gender, VMBO-BL, VMBO-KL, VMBO-GL/TL, HAVO and VWO.

Appendix E Research question 3

Table E1

The main effects of motivation and gender on math achievement within the five tracks. Results from multilevel regression analysis predicting PISA math achievement

Predictor	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	N = 396		*N = 555*		*N = 1172*		*N = 1088*		*N = 1013*	
Student Level	B	p	B	P	β	p	β	p	β	p
Intercept	**-1.36**	**0**	**-0.99**	**0**	**-0.43**	**0**	**0.29**	**0**	**0.83**	**0**
Female	Baseline		Baseline		Baseline		Baseline		Baseline	
Male	**0.22**	**0**	**0.20**	**0**	**0.10**	**0.02**	**0.20**	**0**	0.07	0.10
Self-efficacy	**0.14**	**0**	**0.12**	**0**	**0.18**	**0**	**0.21**	**0**	**0.21**	**0**
Interest value	-0.03	0.39	-0.01	0.91	0.02	0.34	0.03	0.24	**0.11**	**0**
Utility value	0.02	0.62	-0.02	0.64	0.01	0.85	-0.01	0.74	-0.03	0.28

Residual Variance	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	variance	p	Variance	P	variance	p	variance	p	variance	p
Level 2 – intercept	**0.06**	**0.01**	**0.03**	**0.03**	**0.08**	**0**	**0.06**	**0**	**0.07**	**0**
Level 1	**0.2**	**0**	**0.28**	**0**	**0.27**	**0**	**0.26**	**0**	**0.26**	**0**

Note. Results are based on standardized regression coefficients for all variables except for ‘gender’. Female students are used as reference group.

Table E2

The main effects of motivation and gender on math aspirations within the five tracks. Results from multilevel regression analyses predicting PISA math aspirations

Predictor	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	N = 396		*N = 555*		*N = 1172*		*N = 1088*		*N = 1013*	
Student Level	β	p	B	p	β	p	β	p	β	p
Intercept	-0.27	0.10	-0.04	0.70	0.05	0.38	0.10	0.07	**0.11**	**0.04**
Female	Baseline		Baseline		Baseline		Baseline		Baseline	
Male	**0.27**	**0.03**	0.07	0.49	**-0.18**	**0.01**	0	0.97	-0.03	0.66
Self-efficacy	-0.05	0.51	-0.07	0.21	0	0.93	**-0.15**	**0**	-0.06	0.20
Interest value	-0.05	0.49	-0.03	0.61	**-0.12**	**0.01**	**-0.20**	**0**	**-0.23**	**0**
Utility value	**0.37**	**0**	**0.33**	**0**	**0.28**	**0**	**0.25**	**0**	-0.02	0.64

Residual Variance	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	variance	p	variance	p	variance	P	variance	p	variance	p
Level 2 – intercept	-	-	0.06	0.15	0.02	0.27	0.02	0.32	0.02	0.42
Level 1	**0.74**	**0**	**0.77**	**0**	**0.88**	**0**	**1.03**	**0**	**0.90**	**0**

Note. Results are based on standardized regression coefficients for all variables except for 'gender'. Female students are used as reference group.

Table E3

The multiplicative effect of expectancy and value on math achievement within the five tracks. Results from multilevel regression analyses predicting PISA Math achievement

Predictor	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	N = 396		*N = 555*		*N = 1172*		*N = 1088*		*N = 1013*	
Student Level	β	p	β	p	β	p	β	P	β	p
Intercept	**-1.34**	**0**	**-0.98**	**0**	**-0.44**	**0**	**0.30**	**0**	**0.83**	**0**
Female	Baseline		Baseline		Baseline		Baseline		Baseline	
Male	**0.21**	**0**	**0.18**	**0**	**0.10**	**0.02**	**0.20**	**0**	0.07	0.12
Self-efficacy	**0.18**	**0**	**0.12**	**0**	**0.19**	**0**	**0.21**	**0**	**0.21**	**0**
Interest value	**-0.09**	**0.03**	-0.03	0.47	0.02	0.35	0.04	0.20	**0.13**	**0**
Utility value	**0.09**	**0.03**	0.01	0.90	0.01	0.83	-0.01	0.64	-0.05	0.11
Self-efficacy*interest value	**-0.11**	**0.01**	**-0.08**	**0.05**	0.02	0.50	-0.05	0.14	-0.03	0.33
Self-efficacy*utility value	**0.14**	**0**	0.08	0.06	0	1	0.03	0.42	0.04	0.21

Residual Variance	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	variance	p	variance	p	variance	p	variance	P	Variance	p
Level 2 – intercept	**0.06**	**0.01**	**0.03**	**0.02**	**0.08**	**0**	**0.06**	**0**	**0.07**	**0**
Level 1	**0.19**	**0**	**0.27**	**0**	**0.27**	**0**	**0.26**	**0**	**0.26**	**0**

Note. Results are based on standardized regression coefficients for all variables except for 'gender'. Female students are used as the reference group.

Table E4

The multiplicative effect of expectancy and value on math aspirations within the five tracks. Results from multilevel regression analyses predicting PISA math aspirations

Predictor	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	N = 396		*N = 555*		*N = 1172*		*N = 1088*		*N = 1013*	
Student Level	β	P	B	p	β	p	B	p	B	p
Intercept	**-0.25**	**0.02**	0	0.96	0.09	0.10	0.11	0.06	0.09	0.10
Female	Baseline		Baseline		Baseline		Baseline		Baseline	
Male	**0.26**	**0.03**	0.08	0.46	**-0.16**	**0.03**	0	0.99	-0.03	0.66
Self-efficacy	-0.08	0.30	-0.10	0.10	-0.06	0.15	**-0.16**	**0**	-0.08	0.10
Interest value	-0.04	0.65	-0.01	0.89	**-0.12**	**0.01**	**-0.20**	**0**	**-0.29**	**0**
Utility value	**0.33**	**0**	**0.26**	**0**	**0.28**	**0**	**0.26**	**0**	-0.01	0.85
Self-efficacy*interest value	0.01	0.92	0.07	0.34	-0.01	0.91	0.08	0.22	**0.10**	**0.05**
Self-efficacy*utility value	-0.07	0.35	**-0.18**	**0.01**	**-0.15**	**0**	-0.11	0.07	0	0.99

Residual Variance	VMBO-BL		VMBO-KL		VMBO-GL/TL		HAVO		VWO	
	variance	P	Variance	p	variance	p	variance	P	Variance	p
Level 2 – intercept	-	-	0.05	0.20	0.02	0.13	0.02	0.36	0.02	0.47
Level 1	**0.74**	**0**	**0.75**	**0**	**0.83**	**0**	**1.03**	**0**	**0.90**	**0**

Note. Results are based on standardized regression coefficients for all variables except for 'gender'. Female students are used as the reference group.

Table E5

The interaction effects of motivation and tracks on math achievement and aspirations. Results from multilevel regression analysis predicting PISA math achievement and aspirations

	Math achievement		Math aspirations	
Predictor	Model 1		Model 2	
Student Level	β	p	β	P
Intercept	**-1.27**	**0**	-0.08	0.34
Female	Baseline		Baseline	
Male	**0.14**	**0**	-0.03	0.46
VMBO-BL	Baseline		Baseline	
VMBO-KL	**0.35**	**0**	0.11	0.27
VMBO-GL/TL	**0.83**	**0**	0.04	0.61
HAVO	**1.60**	**0**	**0.19**	**0.03**
VWO	**2.05**	**0**	**0.19**	**0.04**
Self-efficacy	**0.14**	**0**	-0.01	0.89
VMBO-BL*self-efficacy	Baseline		Baseline	
VMBO-KL*self-efficacy	-0.01	0.84	-0.04	0.69
VMBO-GL/TL*self-efficacy	0.03	0.46	0	1
HAVO*self-efficacy	0.08	0.10	-0.14	0.13
VWO*self-efficacy	0.06	0.22	-0.05	0.58
Interest value	-0.01	0.82	-0.04	0.63
VMBO-BL*interest value	Baseline		Baseline	
VMBO-KL*interest value	0.02	0.78	0.02	0.88
VMBO-GL/TL*interest value	0.04	0.36	-0.07	0.41
HAVO*interest value	0.04	0.38	-0.16	0.08
VWO*interest value	**0.12**	**0.01**	**-0.19**	**0.04**
Utility value	-0.01	0.84	**0.37**	**0**
VMBO-BL*utility value	Baseline		Baseline	
VMBO-KL*utility value	-0.02	0.73	-0.06	0.55
VMBO-GL/TL*utility value	0	0.09	-0.10	0.25
HAVO*utility value	0	0.98	-0.12	0.18
VWO*utility value	-0.02	0.64	**-0.39**	**0**

Residual variance	Model 1		Model 2	
	variance	p	variance	p
Level 2 – intercept	**0.07**	**0**	0.02	0.07
Level 1	**0.26**	**0**	**0.91**	**0**

Note. VMBO-BL is used as reference group. Results are based on standardized regression coefficients for all variables except for gender, VMBO-BL, VMBO-KL, VMBO-GL/TL, HAVO, and VWO.

Table E6

The interaction effects of expectancy-value and tracks on math achievement and aspirations. Results from multilevel regression analysis predicting PISA math achievement and aspirations

	Math achievement		Math aspirations	
Predictor	Model 1		Model 2	
Student Level	β	p	β	p
Intercept	**-1.26**	**0**	-0.07	0.42
Female	Baseline		Baseline	
Male	**0.13**	**0**	-0.02	0.56
VMBO-BL	Baseline		Baseline	
VMBO-KL	**0.34**	**0**	0.13	0.17
VMBO-GL/TL	**0.81**	**0**	0.09	0.32
HAVO	**1.60**	**0**	**0.19**	**0.04**
VWO	**2.04**	**0**	0.15	0.11
Self-efficacy	**0.17**	**0**	-0.05	0.58
VMBO-BL*self-efficacy	Baseline		Baseline	
VMBO-KL*self-efficacy	-0.04	0.45	-0.03	0.74
VMBO-GL/TL*self-efficacy	0.01	0.83	-0.03	0.78
HAVO*self-efficacy	0.06	0.28	-0.12	0.22
VWO*self-efficacy	0.03	0.59	-0.03	0.72
Interest value	-0.07	0.13	-0.04	0.66
VMBO-BL*interest value	Baseline		Baseline	
VMBO-KL*interest value	0.06	0.37	0.04	0.72
VMBO-GL/TL*interest value	**0.10**	**0.05**	-0.07	0.46
HAVO*interest value	**0.11**	**0.05**	-0.16	0.10
VWO*interest value	**0.20**	**0**	**-0.25**	**0.02**
Utility value	0.07	0.17	**0.34**	**0**
VMBO-BL*utility value	Baseline		Baseline	
VMBO-KL*utility value	-0.07	0.26	-0.10	0.40
VMBO-GL/TL*utility value	-0.07	0.20	-0.07	0.45
HAVO*utility value	-0.08	0.16	-0.08	0.42
VWO*utility value	**-0.12**	**0.04**	**-0.35**	**0**
Self-efficacy*interest value	**-0.11**	**0.02**	-0.02	0.81
VMBO-BL*self-efficacy*interest value	Baseline		Baseline	
VMBO-KL*self-efficacy*interest value	0.04	0.51	0.08	0.47
VMBO-GL/TL*self-efficacy*interest value	**0.13**	**0.02**	0.02	0.86
HAVO*self-efficacy*interest value	0.06	0.30	0.10	0.33
VWO*self-efficacy*interest value	0.08	0.13	0.12	0.22
Self-efficacy*utility value	**0.13**	**0**	-0.05	0.50
VMBO-BL*self-efficacy*utility value	Baseline		Baseline	
VMBO-KL*self-efficacy*utility value	-0.07	0.26	-0.13	0.24
VMBO-GL/TL*self-efficacy*utility value	**-0.13**	**0.01**	-0.10	0.27
HAVO*self-efficacy*utility value	-0.10	0.07	-0.06	0.56

VWO*self-efficacy*utility value	-0.10	0.07	0.05	0.58

Residual variance	Model 1		Model 2	
	variance	p	variance	p
Level 2 – intercept	**0.07**	**0**	**0.02**	**0.05**
Level 1	**0.26**	**0**	**0.89**	**0**

Note. Female students and VMBO-BL is used as reference group. Results are based on standardized regression coefficients for all variables except for gender, VMBO-BL, VMBO-KL, VMBO-GL/TL, HAVO, and VWO.

Appendix F Research question 4

Table F1

The main effects of motivation and tracking on math achievement and aspirations for female and male students separately. Results from multilevel regression analysis predicting PISA math achievement and aspirations

	Math achievement				Math aspirations			
Predictor	Male		Female		Male		Female	
	N = 1425		*N = 1333*		*N = 1385*		*N = 1316*	
Student Level	β	P	B	P	β	p	β	P
Intercept	**-1.10**	**0**	**-1.31**	**0**	**0.01**	**0.93**	**-0.32**	**0.002**
VMBO-BL	Baseline		Baseline		Baseline		Baseline	
VMBO-KL	**0.35**	**0**	**0.38**	**0**	0.03	0.81	**0.24**	**0.05**
VMBO-GL/TL	**0.80**	**0**	**0.86**	**0**	-0.13	0.21	**0.34**	**0**
HAVO	**1.61**	**0**	**1.62**	**0**	0.07	0.52	**0.44**	**0**
VWO	**2.01**	**0**	**2.15**	**0**	0.01	0.96	**0.40**	**0**
Self-efficacy	**0.20**	**0**	**0.16**	**0**	**-0.09**	**0.01**	0	0.91
Interest value	0.01	0.46	**0.07**	**0**	**-0.13**	**0**	**-0.18**	**0**
Utility value	0.01	0.47	-0.04	0.07	**0.20**	**0**	**0.22**	**0**

Residual Variance	Male		Female		Male		Female	
	variance	P	variance	P	variance	p	variance	p
Level 2 – intercept	**0.08**	**0**	**0.05**	**0**	0	0.95	0	0.76
Level 1	**0.27**	**0**	**0.25**	**0**	**0.98**	**0**	**0.93**	**0**

Note. VMBO-BL is used as reference group. Results are based on standardized regression coefficients for all variables except for VMBO-BL, VMBO-KL, VMBO-TL/GL, HAVO, and VWO.

Table F2

The interaction effects of motivation and gender on math achievement and aspirations. Results from multilevel regression analysis predicting PISA math achievement and aspirations

	Math achievement		Math aspirations	
Predictor	Model 1		Model 2	
Student Level	β	p	β	P
Intercept	**-1.25**	**0**	-0.11	0.12
Female	Baseline		Baseline	
Male	**0.14**	**0**	-0.03	0.44
VMBO-BL	Baseline		Baseline	
VMBO-KL	**0.36**	**0**	0.12	0.15
VMBO-GL/TL	**0.80**	**0**	0.09	0.26
HAVO	**1.58**	**0**	**0.23**	**0**
VWO	**2.04**	**0**	**0.18**	**0.03**
Self-efficacy	**0.16**	**0**	0.01	0.72
Female*self-efficacy	Baseline		Baseline	
Male*self-efficacy	0.03	0.19	**-0.11**	**0.01**
Interest value	**0.07**	**0**	**-0.17**	**0**
Female*interest value	Baseline		Baseline	
Male*interest value	**-0.06**	**0.03**	0.05	0.29
Utility value	**-0.04**	**0.05**	**0.22**	**0**
Female*utility value	Baseline		Baseline	
Male*utility value	0.05	0.07	-0.02	0.66

Residual variance	Model 1		Model 2	
	variance	p	variance	p
Level 2 – intercept	**0.07**	**0**	**0.02**	**0.04**
Level 1	**0.26**	**0**	**0.94**	**0**

Note. VMBO-BL and female students are used as reference groups. Results are based on standardized regression coefficients for all variables except for gender, VMBO-BL, VMBO-KL, VMBO-GL/TL, HAVO, and VWO.

Table F3

The multiplicative effect of expectancy and value on math achievement and aspirations for female and male students separately. Results from multilevel regression analyses predicting PISA math achievement and aspirations

	Math achievement				Math aspirations			
Predictor	Male		Female		Male		Female	
	N = 1425		*N = 1333*		*N = 1385*		*N = 1316*	
Student Level	β	p	β	p	β	p	β	P
Intercept	-1.11	0	-1.31	0	0.04	0.70	**-0.26**	**0.01**
VMBO-BL	Baseline		Baseline		Baseline		Baseline	
VMBO-KL	**0.35**	**0**	**0.38**	**0**	0.03	0.79	**0.23**	**0**
VMBO-GL/TL	**0.80**	**0**	**0.86**	**0**	-0.13	0.21	**0.32**	**0**
HAVO	**1.61**	**0**	**1.62**	**0**	0.06	0.55	**0.42**	**0**
VWO	**2.02**	**0**	**2.14**	**0**	0.01	0.93	**0.42**	**0**
Self-efficacy	**0.20**	**0**	**0.17**	**0**	**-0.08**	**0.01**	**-0.09**	**0.02**
Interest value	0.01	0.71	**0.07**	**0**	**-0.13**	**0**	**-0.17**	**0**
Utility value	0.01	0.54	-0.03	0.13	**0.23**	**0**	**0.18**	**0**
Self-efficacy*interest value	0.01	0.78	-0.02	0.37	0.04	0.30	0.01	0.76
Self-efficacy*utility value	0.01	0.65	0.03	0.12	**-0.11**	**0**	**-0.19**	**0**

Residual Variance	Male		Female		Male		Female	
	variance	p	Variance	p	variance	p	variance	P
Level 2 – intercept	**0.08**	**0**	**0.05**	**0**	0	0.86	0	0.85
Level 1	**0.27**	**0**	**0.25**	**0**	**0.96**	**0**	**0.90**	**0**

Note. VMBO-BL is used as reference group. Results are based on standardized regression coefficients for all variables except for VMBO-BL, VMBO-KL, VMBO-GL/TL, HAVO, and VWO.

Table F4

The interaction effects of expectancy-value and gender on math achievement and aspirations. Results from multilevel regression analysis predicting PISA math achievement and aspirations

	Math achievement		Math aspirations	
Predictor	Model 1		Model 2	
Student Level	β	p	B	P
Intercept	**-1.26**	**0**	-0.06	0.47
Female	Baseline		Baseline	
Male	**0.13**	**0**	**-0.05**	**0.26**
VMBO-BL	Baseline		Baseline	
VMBO-KL	**0.35**	**0**	0.14	0.16
VMBO-GL/TL	**0.81**	**0**	0.09	0.29
HAVO	**1.60**	**0**	**0.20**	**0.03**
VWO	**2.04**	**0**	0.16	0.09
Self-efficacy	**0.16**	**0**	**-0.05**	**0.56**
VMBO-BL*self-efficacy	Baseline		Baseline	
VMBO-KL*self-efficacy	-0.04	0.77	-0.04	0.71
VMBO-GL/TL*self-efficacy	0.04	0.41	-0.02	0.81
HAVO*self-efficacy	0.08	0.10	-0.11	0.22
VWO*self-efficacy	0.06	0.25	-0.03	0.79
Female*Self-efficacy	Baseline		Baseline	
Male*Self-efficacy	0.02	0.39	-0.01	0.75
Interest value	-0.05	0.38	-0.03	0.75
VMBO-BL*interest value	Baseline		Baseline	
VMBO-KL*interest value	-0.01	0.89	0.03	0.81
VMBO-GL/TL*interest value	0.03	0.50	-0.08	0.41
HAVO*interest value	0.04	0.50	-0.17	0.08
VWO*interest value	**0.12**	**0.02**	**-0.26**	**0.02**
Female*interest value	Baseline		Baseline	
Male*interest value	-0.03	0.26	0	0.98
Utility value	0.04	0.39	**0.32**	**0**
VMBO-BL*utility value	Baseline		Baseline	
VMBO-KL*utility value	0	0.95	-0.09	0.42
VMBO-GL/TL*utility value	0.01	0.80	-0.07	0.50
HAVO*utility value	0.01	0.92	-0.08	0.45
VWO*utility value	-0.03	0.57	**-0.34**	**0**
Female*utility value	Baseline		Baseline	
Male*utility value	0.03	0.30	0.02	0.70
Self-efficacy*interest value	0	0.91	-0.01	0.90
VMBO-BL*self-efficacy*interest value	Baseline		Baseline	
VMBO-KL*self-efficacy*interest value	-0.07	0.15	0.06	0.59
VMBO-GL/TL*self-efficacy*interest value	0.02	0.63	0	0.97

HAVO*self-efficacy*interest value	-0.05	0.22	0.08	0.47
VWO*self-efficacy*interest value	-0.03	0.49	0.10	0.33
Female*self-efficacy*interest value	Baseline		Baseline	
Male*self-efficacy*interest value	0.02	0.60	0.02	0.65
Self-efficacy*utility value	**0.14**	**0**	-0.09	0.30
VMBO-BL*self-efficacy*utility value	Baseline		Baseline	
VMBO-KL*self-efficacy*utility value	-0.06	0.31	-0.12	0.28
VMBO-GL/TL*self-efficacy*utility value	**-0.13**	**0.01**	-0.10	0.27
HAVO*self-efficacy*utility value	-0.10	0.07	-0.05	0.59
VWO*self-efficacy*utility value	-0.09	0.08	0.05	0.63
Female*self-efficacy*utility value	Baseline		Baseline	
Male*self-efficacy*utility value	-0.03	0.31	0.06	0.24

Residual variance	Model 1		Model 1	
	variance	p	variance	p
Level 2 – intercept	**0.07**	**0**	**0.02**	**0.05**
Level 1	**0.26**	**0**	**0.89**	**0**

Note. VMBO-BL and female students are used as reference groups. Results are based on standardized regression coefficients for all variables except for gender, VMBO-BL, VMBO-KL, VMBO-GL/TL, HAVO, and VWO.

Appendix G Figures multiplicative effects

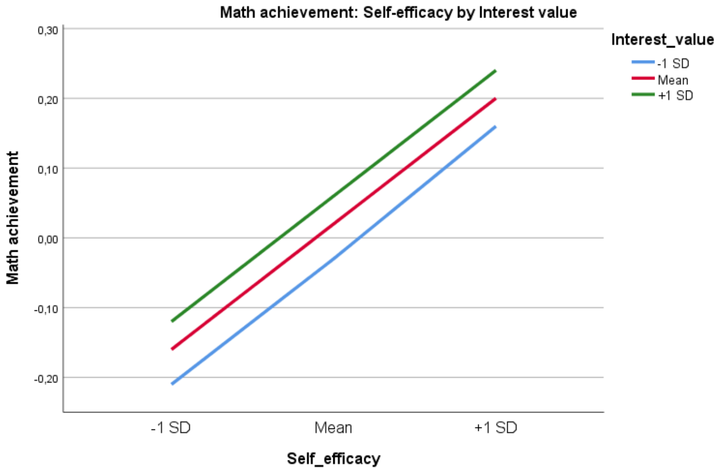

Figure G1

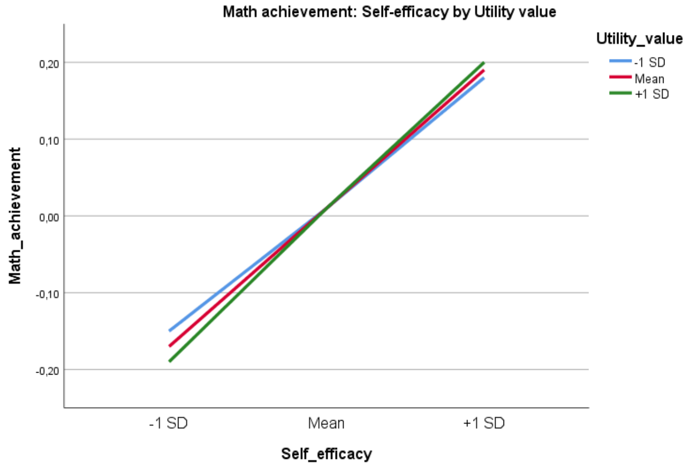

Figure G2

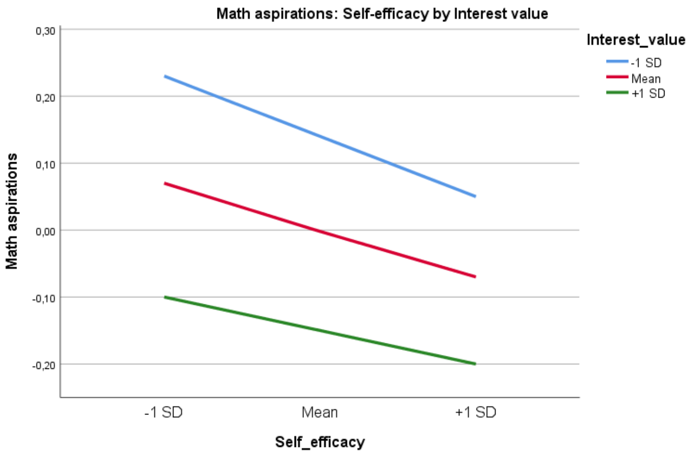

Figure G3

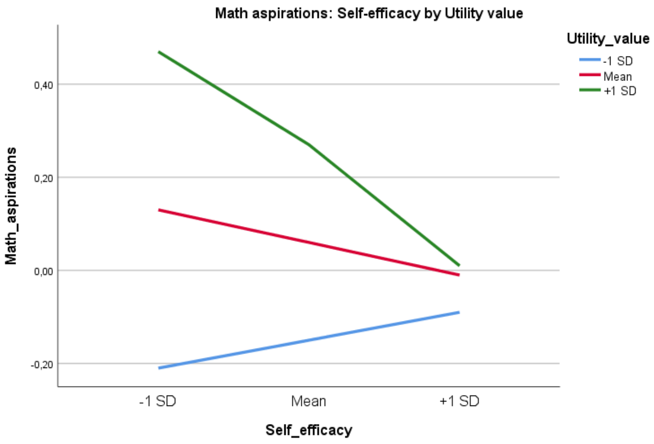

Figure G4